

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:10:59 ; Search time 128.03 Seconds  
(without alignments)  
985.003 Million cell updates/sec

Title: US-09-235-416-1

Perfect score: 4030

Sequence: 1 MSGGNKVVVRFPNARE.....ELRQAOQMEALAKTAKQEF 784

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*  
1: /cgn2\_6/ptodata/2/paa/US06\_COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/paa/US07\_COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/paa/US08\_COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pcp.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pcp.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pcp.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pcp.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pcp.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pcp.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pcp.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pcp.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pcp.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pcp.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pcp.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pcp.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pcp.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pcp.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pcp.\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pcp.\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pcp.\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pcp.\*  
24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	4030	100.0	784	PCT-US99-01355-1	Sequence 1, Appli
2	4030	100.0	784	US-09-235-416-1	Sequence 1, Appli
3	4030	100.0	784	US-09-654-850-1	Sequence 1, Appli
4	2054	51.0	533	US-09-417-507-37298	Sequence 37298, A
5	1691	42.0	1153	US-09-718-563-2	Sequence 2, Appli
6	1691	42.0	1153	US-09-718-804-2	Sequence 2, Appli
7	1691	42.0	1153	US-09-718-842-2	Sequence 2, Appli
8	1680.5	41.7	1770	US-09-718-563-6	Sequence 6, Appli
9	1680.5	41.7	1770	US-09-718-804-6	Sequence 6, Appli
10	1680.5	41.7	1770	US-09-718-842-6	Sequence 6, Appli

11	1669.5	41.4	1793	US-60-258-275-444	Sequence 444, App
12	1668.5	41.4	1816	US-09-718-563-8	Sequence 8, Appli
13	1668.5	41.4	1816	US-09-718-804-8	Sequence 8, Appli
14	1668.5	41.4	1816	US-09-718-842-8	Sequence 8, Appli
15	1668.5	41.4	1816	US-60-130-217-6	Sequence 6, Appli
16	1660.5	41.2	893	US-09-488-725A-3179	Sequence 3179, Ap
17	1658	41.1	1103	US-09-467-946-1	Sequence 1, Appli
18	1635	40.6	1773	US-60-167-217-18615	Sequence 18615, A
19	1635	40.6	1773	US-60-173-464-15257	Sequence 15257, A
20	1635	40.6	1773	US-60-191-637-18572	Sequence 18572, A
21	1635	40.6	1773	US-60-191-681-14693	Sequence 14693, A
22	1431.5	35.5	1362	US-09-380-828-7631	Sequence 7631, Ap
23	1386.5	34.7	1931	US-60-167-217-15727	Sequence 15727, A
24	1386.5	34.7	1931	US-60-173-464-12890	Sequence 12890, A
25	1386.5	34.7	1931	US-60-191-637-15722	Sequence 15722, A
26	1386.5	34.7	1931	US-60-191-681-12455	Sequence 12455, A
27	1386.5	34.7	1931	US-09-721-689-2	Sequence 2, Appli
28	1311.5	32.5	1375	US-09-721-832-2	Sequence 2, Appli
29	1311.5	32.5	1375	US-09-722-139-2	Sequence 2, Appli
30	1311.5	32.5	1375	US-09-722-139-2	Sequence 2, Appli
31	1278.5	31.7	504	PCT-US00-09066-115	Sequence 115, App
32	1276	31.7	503	PCT-US00-09066-116	Sequence 116, App
33	1251	31.0	421	US-09-488-725A-6751	Sequence 6751, Ap
34	1133	28.1	348	US-09-718-563-4	Sequence 4, Appli
35	1133	28.1	348	US-09-718-804-4	Sequence 4, Appli
36	1133	28.1	348	US-09-718-842-4	Sequence 4, Appli
37	1129	28.0	354	US-09-718-563-10	Sequence 10, Appli
38	1129	28.0	354	US-09-718-804-10	Sequence 10, Appli
39	1129	28.0	354	US-09-718-842-10	Sequence 10, Appli
40	1128	28.0	1174	US-60-173-464-1013	Sequence 1013, A
41	1128	28.0	1174	US-60-173-464-9672	Sequence 9672, A
42	1128	28.0	1174	US-60-191-637-11636	Sequence 11636, A
43	1128	28.0	1174	US-60-191-681-9368	Sequence 9368, Ap
44	1072	25.6	352	US-09-580-828-4	Sequence 4, Appli
45	1035	25.7	1121	US-09-619-049-1524	Sequence 1524, Ap

ALIGNMENTS

RESULT 1  
PCT-US99-01355-1  
; Sequence 1, Application PCT/US9901355  
; GENERAL INFORMATION:  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Goldstein, Lawrence S. B.  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor  
; FILE REFERENCE: 18557C-000710PC  
; CURRENT APPLICATION NUMBER: PCT/US99/01355  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER FILING DATE: 1999-01-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 784  
; TYPE: PRT  
; ORGANISM: Thermomyces lanuginosus  
; FEATURE:  
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed  
; OTHER INFORMATION: microtubule motor protein  
; NAME/KEY: DOMAIN  
; LOCATION: (1) (357)  
; OTHER INFORMATION: kinesin-like microtubule motor domain  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (358)..(442)  
; OTHER INFORMATION: neck domain links motor domain to stalk domain  
; FEATURE:  
; NAME/KEY: DOMAIN

LOCATION: (443)..(601)  
OTHER INFORMATION: stalk domain, unc-104 family domain  
NAME/KEY: DOMAIN  
LOCATION: (602)..(784)  
OTHER INFORMATION: tail domain  
NAME/KEY: VARIANT  
LOCATION: (713)  
OTHER INFORMATION: polymorphic variant #1 Val -> Ile  
NAME/KEY: VARIANT  
LOCATION: (762)  
OTHER INFORMATION: polymorphic variant #2 Asp -> Glu  
NAME/KEY: VARIANT  
LOCATION: (774)  
OTHER INFORMATION: polymorphic variant #3 Glu -> Asp  
PCT-US99-01355-1

Query Match 100.0%; Score 4030; DB 1; Length 784;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGNNIKVYVRPNAREIDRGAKCIVRMGNTILTPPGAEEKARKSGKTIMDGP 60  
DB 1 MSGGNNIKVYVRPNAREIDRGAKCIVRMGNTILTPPGAEEKARKSGKTIMDGP 60

QY 61 APAFDRSYWSDKNAFYARQEDLFDGLVPLDPAFGYNKNCIFAYGQTGSGKSYSMWG 120  
DB 61 APAFDRSYWSDKNAFYARQEDLFDGLVPLDPAFGYNKNCIFAYGQTGSGKSYSMWG 120

QY 121 YGKEHGVIPRICQDMFRRINELQKDNLTCTVEYSYLEIYNERVDLLNPNSTKGNLKYRE 180  
DB 121 YGKEHGVIPRICQDMFRRINELQKDNLTCTVEYSYLEIYNERVDLLNPNSTKGNLKYRE 180

QY 181 HPSTGTPVEDLAKLVYRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTITQKWH 240  
DB 181 HPSTGTPVEDLAKLVYRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTITQKWH 240

QY 241 DEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLSTLGRVTAALADMS 300  
DB 241 DEETKMDTEKVAKISLVLAGSERATSTGATGARLKEGAEINRSLSTLGRVTAALADMS 300

QY 301 GKQKNQVLPYRDSVLTWLLKDSLGNSKMTAMTALSPADINFEETSLTRYADSAKRK 360  
DB 301 GKQKNQVLPYRDSVLTWLLKDSLGNSKMTAMTALSPADINFEETSLTRYADSAKRK 360

QY 361 NHAVVWEDNARMTRLEKPELQALRSKQSGGGGAGGSGGVPVEYPTDPLEKQIV 420  
DB 361 NHAVVWEDNARMTRLEKPELQALRSKQSGGGGAGGSGGVPVEYPTDPLEKQIV 420

QY 421 STQOPDATVKKMSKAEIVPOLNOSKLYRDLNQTWEKELAKTEETHKEREALEELGIS 480  
DB 421 STQOPDATVKKMSKAEIVPOLNOSKLYRDLNQTWEKELAKTEETHKEREALEELGIS 480

QY 481 EKGFWGPTBSKMPHLVNLSDPLAELCYVNIKPGQTRGVNVDQTAQELRLNGSKLK 540  
DB 481 EKGFWGPTBSKMPHLVNLSDPLAELCYVNIKPGQTRGVNVDQTAQELRLNGSKLK 540

QY 541 EHCFTFNDVNTVIVPNEKAAVWVNGVRIDKPTLRSGYRILILGDHFIFREHNPEARAE 600  
DB 541 EHCFTFNDVNTVIVPNEKAAVWVNGVRIDKPTLRSGYRILILGDHFIFREHNPEARAE 600

QY 601 ROEQSLLRHSVTNSQLGSPAPGRDHDTLSKAGSDADGDSRSDPLPHFRCKDSQWYARR 660  
DB 601 ROEQSLLRHSVTNSQLGSPAPGRDHDTLSKAGSDADGDSRSDPLPHFRCKDSQWYARR 660

QY 661 EAAASAILGLDQKISHLTDELDALFDYQKARAVRGLVEDNEDSDSSQSPFVRDKYMSN 720  
DB 661 EAAASAILGLDQKISHLTDELDALFDYQKARAVRGLVEDNEDSDSSQSPFVRDKYMSN 720

QY 721 GTIDNESLDTAITMPCPTRSDDDGALFFGDKKSODASNVVDVEELRQQAQMEALKTA 780  
DB 721 GTIDNESLDTAITMPCPTRSDDDGALFFGDKKSODASNVVDVEELRQQAQMEALKTA 780

QY 781 KQEF 784  
DB 781 KQEF 784

RESULT 2  
US-09-235-416-1  
Sequence 1, Application US/09235416A  
GENERAL INFORMATION:  
APPLICANT: Sakowicz, Roman  
APPLICANT: Goldstein, Lawrence S. B.  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor  
FILE REFERENCE: 18557C-000710US  
CURRENT APPLICATION NUMBER: US/09/235,416A  
CURRENT FILING DATE: 1999-01-22  
EARLIER APPLICATION NUMBER: WO PCT/US99/01355  
EARLIER FILING DATE: 1999-01-22  
EARLIER APPLICATION NUMBER: US 60/072,361  
EARLIER FILING DATE: 1998-01-23  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 784  
TYPE: PRT  
ORGANISM: Thermomyces lanuginosus  
FEATURE:  
OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed  
OTHER INFORMATION: microtubule motor protein  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (1)..(357)  
OTHER INFORMATION: kinesin-like microtubule motor domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (358)..(442)  
OTHER INFORMATION: neck domain links motor domain to stalk domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (443)..(601)  
OTHER INFORMATION: stalk domain, unc-104 family domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (602)..(784)  
OTHER INFORMATION: tail domain  
US-09-235-416-1

Query Match 100.0%; Score 4030; DB 1; Length 784;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGNNIKVYVRPNAREIDRGAKCIVRMGNTILTPPGAEEKARKSGKTIMDGP 60  
DB 1 MSGGNNIKVYVRPNAREIDRGAKCIVRMGNTILTPPGAEEKARKSGKTIMDGP 60

QY 61 APAFDRSYWSDKNAFYARQEDLFDGLVPLDPAFGYNKNCIFAYGQTGSGKSYSMWG 120  
DB 61 APAFDRSYWSDKNAFYARQEDLFDGLVPLDPAFGYNKNCIFAYGQTGSGKSYSMWG 120

QY 121 YGKEHGVIPRICQDMFRRINELQKDNLTCTVEYSYLEIYNERVDLLNPNSTKGNLKYRE 180  
DB 121 YGKEHGVIPRICQDMFRRINELQKDNLTCTVEYSYLEIYNERVDLLNPNSTKGNLKYRE 180

QY 181 HPSTGTPVEDLAKLVYRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTITQKWH 240  
DB 181 HPSTGTPVEDLAKLVYRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTITQKWH 240

QY	241	DEETKNDTEKVAKISLVLDAGSERATSGATGACRLKEGAENRSITLGRVIALADNMSS	300
	241	DEETKNDTEKVAKISLVLDAGSERATSGATGACRLKEGAENRSITLGRVIALADNMSS	300
QY	301	GKQKKNQLVPYRDSVLTWLLKDSLGNSTMATAISPAIDNFEETLSTLRVADSNAKRIK	360
QY	301	GKQKKNQLVPYRDSVLTWLLKDSLGNSTMATAISPAIDNFEETLSTLRVADSNAKRIK	360
QY	361	NHAYVEDPNARIRELKEELAEOLRSKLQSGGGGGAGGSGPVVEESYPPDTPLEKQIV	420
QY	361	NHAYVEDPNARIRELKEELAEOLRSKLQSGGGGGAGGSGPVVEESYPPDTPLEKQIV	420
QY	421	STQODPATYKKNKAEIVQOLNQSSEKLYRDLNQTWEELAKTEERKEPEANLEELGTSI	480
QY	421	STQODPATYKKNKAEIVQOLNQSSEKLYRDLNQTWEELAKTEERKEPEANLEELGTSI	480
QY	481	EKGFVGPYHRSKEMPLVNLSDPPLAECLVYNIKPQOTRGVYNQDQOAEIRLNGSLTK	540
QY	481	EKGFVGPYHRSKEMPLVNLSDPPLAECLVYNIKPQOTRGVYNQDQOAEIRLNGSLTK	540
QY	541	EICTFENDVNWTVIPNEKAAVWNGVRIDKPTRLKSGYKRIILGDHFIFRNPHEARAE	600
QY	541	EICTFENDVNWTVIPNEKAAVWNGVRIDKPTRLKSGYKRIILGDHFIFRNPHEARAE	600
QY	601	ROQSILRHSVYNSQLSGPACPHDRTLSKAGSDADGSDSDPLPIPHFGDSDWFTYARR	660
QY	601	ROQSILRHSVYNSQLSGPACPHDRTLSKAGSDADGSDSDPLPIPHFGDSDWFTYARR	660
QY	661	BAASA1LGLDQK1SHLTDDELDALFDVQOKARVRVRLGVLEDNEDSDSQSPVPRQKYM	720
QY	661	BAASA1LGLDQK1SHLTDDELDALFDVQOKARVRVRLGVLEDNEDSDSQSPVPRQKYM	720
QY	721	GTIDNPSLDTA1TMPGTPRSDDDGALFFGDKSKQDASNVQVDELRQOQAQMEALKTA	780
QY	721	GTIDNPSLDTA1TMPGTPRSDDDGALFFGDKSKQDASNVQVDELRQOQAQMEALKTA	780
QY	781	KQEF 784	
QY	781	KQEF 784	
QY	781	KQEF 784	

```

RESULT      3
US-09-654-850-1
: Sequence 1, Application US/09654850
: GENERAL INFORMATION:
: APPLICANT: Sakowicz, Roman
: APPLICANT: Goldstein, Lawrence S. B.
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
: TITLE OF INVENTION: Protein
: FILE REFERENCE: 18557C-000710US
: CURRENT APPLICATION NUMBER: US/09/654,850
: PRIOR FILING DATE: 2000-09-05
: PRIOR APPLICATION NUMBER: US/09/235,416
: PRIOR FILING DATE: 1998-01-22
: PRIOR APPLICATION NUMBER: US 60/072,361
: PRIOR FILING DATE: 1998-01-23
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 784
: TYPE: PRT
: FEATURE:
: OTHER INFORMATION: Thermomyces lanuginosus
: FEATURE:
: OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
: OTHER INFORMATION: microtubule motor protein
: NAME/KEY: DOMAIN357
: LOCATION: 357
: OTHER INFORMATION: kinesin-like microtubule motor domain
: NAME/KEY: DOMAIN
: LOCATION: (358) .. (442)
: OTHER INFORMATION: neck domain links motor domain to stalk domain
: NAME/KEY: DOMAIN

```

```

: LOCATION: (443)..(601)
: OTHER INFORMATION: stalk domain, unc-104 family domain
: NAME/SET: DOMAIN
: LOCATION: (602)..(784)
: OTHER INFORMATION: tail domain
US-09-654-450-1

Query Watch      100.0%: Score 4030: DB 20: Length 784;
Best Local Similarity
Matches 784: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 MSGGNGIKVVRVRFNAREIDRGACACIVRMEGNQTTILTPPGAEBEAKRSKGTIDMGPK 60
   |||||
Db 1 MSGGNGIKVVRVRFNAREIDRGACACIVRMEGNQTTILTPPGAEBEAKRSKGTIDMGPK 60

QY 61 AFAPDRSWSFKDNAPYARQEDLFODLGVPLLDNAFKYNNCIFAYGQTGSKSYSNMG 120
   |||||
Db 61 AFAPDRSWSFKDNAPYARQEDLFODLGVPLLDNAFKYNNCIFAYGQTGSKSYSNMG 120

QY 121 YKEHGVTPIRQCDMFRINELOKDKNCTCTVENSYLEIYMERVROLNLPSTKGNLKVRE 180
   |||||
Db 121 YKEHGVTPIRQCDMFRINELOKDKNCTCTVENSYLEIYMERVROLNLPSTKGNLKVRE 180

QY 181 PHSCTGCVPILOAKLVRSFOEINLMDGKARKVAATNNKTSRSSHAFVLTITOKWH 240
   |||||
Db 181 PHSCTGCVPILOAKLVRSFOEINLMDGKARKVAATNNKTSRSSHAFVLTITOKWH 240

QY 241 DEETKNDYTKVAKISLVDLAGSERATSTGATGARLKGEANRSLSTGLRVTAALADMS 300
   |||||
Db 241 DEETKNDYTKVAKISLVDLAGSERATSTGATGARLKGEANRSLSTGLRVTAALADMS 300

QY 301 GKQKKNLQVPRDSYLTWLLKDSLGSNSWTAMTAISPADINFEETLSTLRVADSARKIK 360
   |||||
Db 301 GKQKKNLQVPRDSYLTWLLKDSLGSNSWTAMTAISPADINFEETLSTLRVADSARKIK 360

QY 361 NHAVYEDPNARMIRELKAEALQRLSKQSGGGGAGSGGVPEESTPDTPLEKQIV 420
   |||||
Db 361 NHAVYEDPNARMIRELKAEALQRLSKQSGGGGAGSGGVPEESTPDTPLEKQIV 420

QY 421 SIQPDATVTKKSKAEATVQLNQSEKLYRDLNQTWEKLANTEBIHKREPAALBELGISI 480
   |||||
Db 421 SIQPDATVTKKSKAEATVQLNQSEKLYRDLNQTWEKLANTEBIHKREPAALBELGISI 480

QY 481 EKGFGPGVHSEKMPHLYNSDDPLLAELGLVYNIKPGQTVGVNWDQTOEIRLNGSKILK 540
   |||||
Db 481 EKGFGPGVHSEKMPHLYNSDDPLLAELGLVYNIKPGQTVGVNWDQTOEIRLNGSKILK 540

QY 541 ECHTFCNDYDNYTVIPNEKAAYVNCVRIDKPTRLRSYGRVILGDPIFRFNHPEEARAE 600
   |||||
Db 541 ECHTFCNDYDNYTVIPNEKAAYVNCVRIDKPTRLRSYGRVILGDPIFRFNHPEEARAE 600

QY 601 RQEGSLLRHYSNQLGSPAPGRHRTLSKAGSDAGDSRSDSPPLPHFGKOSDMWFYAR 660
   |||||
Db 601 RQEGSLLRHYSNQLGSPAPGRHRTLSKAGSDAGDSRSDSPPLPHFGKOSDMWFYAR 660

QY 661 EASALITGLDQKISHLTDDDELQALFDVQKARVRCGLVEDNESDSSGSPFRDXYMSN 720
   |||||
Db 661 EASALITGLDQKISHLTDDDELQALFDVQKARVRCGLVEDNESDSSGSPFRDXYMSN 720

QY 721 CTDINFSLDTATWMPGTDRSDDDGALPFQDKSKQDASNVDVEELRQOQAWEEALKTA 780
   |||||
Db 721 CTDINFSLDTATWMPGTDRSDDDGALPFQDKSKQDASNVDVEELRQOQAWEEALKTA 780

QY 781 KQEF 784
   |||||
Db 781 KQEF 784

```

RESULT 4  
US-09-417-507-37298  
; Sequence 37298, Application US/09417507  
; GENERAL INFORMATION:

APPLICANT: KEITH G. WEINSTOCK ET AL.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS  
FILE OF INVENTION: FUMIGATUS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PATH99-10  
CURRENT APPLICATION NUMBER: US/09/417,507  
CURRENT FILING DATE: 1999-10-14  
NUMBER OF SEQ ID NOS: 44312  
SEQ ID NO 37298  
LENGTH: 533  
TYPE: PRT  
ORGANISM: A.fumigatus  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (515)  
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

Query Match 51.0%; Score 2054; DB 18; Length 533;  
Best Local Similarity 77.0%; Pred. No. 7,1e-153;  
Matches 408; Conservative 46; Mismatches 68; Indels 8; Gaps 5;

QY 20 EIDSGAKCIIVMEGNOTILTPPGABEAKARKS-GTTIMDGPKAFAFORSYFQKNAPNY 78  
DB 8 EIDRGAKCIVOMKGNOTILVPPGADKSRKAGGAVGEGPKTFAFORSYFQKNAPNY 67  
QY 79 ARQEDLFQDLGVPILLONAFKYNCCIFAYGOTGSGKYSMMGYKGEHGVIPRIQDNFRR 138  
DB 68 AGQDNLFADLGVPLLONAFQYNNCCIFAYGOTGSGKYSMMGYKGEHGVIPRIQDNFRR 127  
QY 139 INELQDKNIACTVEYSLEYTYNERNVROLNPNSTGNLYRHPHSTGTYVEDLAKLVRS 198  
DB 128 TAKWEDKLNIACTVEYSLEYTYNERNVROLNPNSTGNLYRHPHSTGTYVEDLAKLVRS 187  
QY 199 FOETENLMDGNKARTVAATNMNNTSSSHAVFTLTITQKHDEKMDTKVAKISLVD 258  
DB 188 FEEDHLMMDGNKARTVAATNMNNTSSSHAVFTLTITQKHDEKMDTKVAKISLVD 247  
QY 259 LAGSERATSTGATGARKEGAEINSLSTLGRVIAALADMSSGKQKNQLVPRDSVLTM 318  
DB 248 LAGSERANSTGATGARKEGAEINSLSTLGRVIAALADMSSGKQKNQNVPRDSVLTM 307  
QY 319 LKDSLGNSMTAMIAAISPADINFEETLTSLRYADSARKIKNHAVNEDPNARMIRELK 378  
DB 308 LKDSLGNSMTAMIAAISPADINFEETLTSLRYADSARKIKNHAVNEDPNARMIRELK 367  
QY 379 BELAQURKLSGGGGGGGGAGSGG-PVRESYPPDPLEKQIVSTQQPDATVKKMSKAEI 437  
DB 368 BELAQURKLSGGGGGGGGAGSGG-PVRESYPPDPLEKQIVSTQQPDATVKKMSKAEI 423  
QY 438 VEQUNSEKLYRDJNMTWEEKL-AKTEBIHKERAEALLEGISIEKFGVGYHSEKMPHL 496  
DB 424 VEQUNSEKLYRDJNMTWEEKL-AKTEBIHKERAEALLEGISIEKFGVGYHSEKMPHL 483  
QY 497 VNLSDPPLAELVYIKPGQTVGVNMQDTQAEIRLM-GSKILKEHCFTF 545  
DB 484 VNLSDPPLAELVYIKPGQTVGVNMQDTQAEIRLM-GSKILKEHCFTF 533

RESULT 5  
US-09-718-563-2  
GENERAL INFORMATION: US/09718563  
APPLICANT: Beraud, Christophe  
TITLE OF INVENTION: Novel motor proteins and methods for  
FILE REFERENCE: Their use  
CURRENT APPLICATION NUMBER: US/09/718,563  
CURRENT FILING DATE: 2000-11-22  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2

LENGTH: 1153  
TYPE: PRT  
ORGANISM: Human  
US-09-718-563-2

Query Match 42.0%; Score 1691; DB 21; Length 1153;  
Best Local Similarity 43.6%; Pred. No. 9,6e-124;  
Matches 382; Conservative 137; Mismatches 235; Indels 122; Gaps 19;

QY 4 GGNLKVVRVVRPNARETRDCAKCTVRMEGNOTILTPPGABEAKARKSGKTIMGPKAKA 63  
DB 3 GASVGVAVRVRPNARETRDCAKCTVRMEGNOTILTPPGABEAKARKSGKTIMGPKAKA 51  
QY 64 FDRSYNSF-DKNAPNTAFQEDFDLGVPLLDNAPKYNCCIFAYGOTGSGKYSMMGYG 122  
DB 52 FDRSYNSHTSPEDCFASONRVYNDIGKEMLLHAFEGYNYCIFYAYGOTGAGKSYTMGKQ 111  
QY 123 KEH--GVIPRIQDMFERNLQKDNLTCTVEYSLEYTYNERNVROLNPNSTGNLKVRE 180  
DB 112 BESQAGIIPOLCELPPEKIND-NCNEEKSYSVEYSMEIYCYRVRVRLNPNKGNLKVRE 170  
QY 181 HPSTGTYVEDLAKLVRSFORIENLMDGNKARTVAATNMNNTSSSHAVFTLTITQKH 240  
DB 171 HPLLGPVVEDLSKLAVTSYTDIADLDAGNKNARTVAATNMNNTSSSHAVFTVFTOKKH 230  
QY 241 DEETKMDTEKAKISLVDLAGSERATSTGATGARKEGAEINSLSTLGRVIAALADMSS 300  
DB 231 DNETNLTSEKYSISLVDLAGSERADSTGATGTRLKEGANKINSLTTLGRVIAALAEVSK 290  
QY 301 GKQKKNQLVPRDSVLTMKDSLGNSMTAMIAAISPADINFEETLTSLRYADSARKIK 360  
DB 291 -KKKTKDTFTPYRDSVLTMKDSLGNSMTAMIAAISPADINFEETLTSLRYADSARKIK 349  
QY 361 NHAVNEDPNARMIRELKEELAQRLSKLQSGGGGGGAGSGGPPVEESTYPPDPLEKQ-- 418  
DB 350 CNVAINEDPNARMIRELKEELAEVTRIKDLJLRAGIGDIIDTSMGSLT--SSPSSCLSSQVG 407  
QY 419 ---IVSIQPDATVKKMSKAEIYEOLNOSKLYRDLNOTWEEKLAKTEBIHKERAEAL 475  
DB 408 LTVSTSIQ--ERIMSTPGGEEAERLKESEKIIITAEINLWEEKLEKTEATRMEREALAE 465  
QY 476 LGSIEK--GFVGPYHSEKMPHLVNI-SDDPLAELVYIKPGQTVGVNMQDTQAEIRL 533  
DB 466 MGVAIREDDGGTGLGVSPKKTPLHVNLEDPLMSECLLYIKDGTIRVGQADAEERDIDL 525  
QY 534 NGSKILKEHCFTFENV---DNVVTIVPNEKAAMVWGVNRIDKPTPLRSYRILGDFH 588  
DB 526 SGAIHKEEHCIFERSNSGEVITLPECEKSETYVNGKVSQPLVNSGRNIRINGKNV 585  
QY 589 FRFHPEEARAER-----QESLRLHSHVNSQ-----LGSF 619  
DB 586 FRFHPEEARAERKTPSNATPEPVDVWTFQAQRELLERKQIGDKMKEKLEMEILYKK 645  
QY 620 APGRHDTLUSKAGSDADGSDRSDS----- 643  
DB 646 EKEEADLLLEQRLDADSDGSDSKRSCEESWKLTSLREKLPPSKLQTVYKKGGLUPS 705  
QY 644 -----PLPHFR--GKSDWFFVARRAASAILGDDOKISHLTD-----DELDALFD 686  
DB 706 GKKREPTKMYQIPORRLRSKQKWTISDLKIQVKEICTEVA-LNDFRHSQREIATAI 764  
QY 687 DVQKARVARGLVEDNEUSDQSSFPVRDYKYNNGTIDNFSLDTAITNPGTFRSD-- 742  
DB 765 VKMELCAMYKGDQPNP-RDSNRAV-ARDVWDTVGVGDEKIEDVATGKSTGDTVDLLKVH 822  
QY 743 -DGDALFFDCKKQDASVVDYELRQQAQMEAL 777  
DB 823 DTKLEDILQEVKQNNKRDDEIRVLANKMLMEKVL 858

RESULT 6  
US-09-718-804-2



; Sequence 2, Application US/09718804  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; FILE REFERENCE: 1021  
 ; CURRENT APPLICATION NUMBER: US/09/718 804  
 ; CURRENT FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1153  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-718-804-2

Query Match 42.0%; Score 1691; DB 21; Length 1153;  
 Best Local Similarity 43.6%; Pred. No. 9,6e-124;  
 Matches 382; Conservative 137; Mismatches 235; Indels 122; Gaps 19;

Qy 4 GGNIKVVRVPFNAREIDRCAKIVRMENQOTILPPPGAEEKARKSKGTINDGPKAFA 63  
 Db 3 GASVAVVVRPFPNSBETSKESKCIIOQNGNSTSIINPNKE-----APKFS 51  
 Qy 64 FDRSYNSF-DKNAPYARQEDFDLGVPLDNAPKNCIFAYGOTGSKSYSMGTYG 122  
 Db 52 FDSYNSHTSPEDPCFASQNRVYNDIGKMLLHAFEGYVCIFAYGOTGAGKSYTMGQ 111  
 Qy 123 KEH--GVIPRICQDMFRINELQDKMLTCVVEVSYLETYNERYVDLLNPSTKGNLKVRE 180  
 Db 112 EESQAGIIPOLCELEFKIND-NCNEEMSYSEVSYMEYTCERYVDLLNPKNGLRVRE 170  
 Qy 181 HPSTGPIYVEDLAKLVRSFQENLDGNCARIVTAATNMNTHSSRHAVFTLTQKWH 240  
 Db 171 HPLGPIYVEDLSKLVSTYDIALMDAGNKARTVAATNMNTHSSRHAVFTIVTQKH 230  
 Qy 241 DEETKMDTKVAKISLVDLAGESEATSGATGAKEGAEINRSLSLTGLRVIAALADMS 300  
 Db 231 DNETNLSTKVSISLVDLAGESEATSGATGAKEGAEINRSLSLTGLKVISALAEVSK 290  
 Qy 301 GKOKNOLVPRDSVLTLMLKDSLGGSNTAMIAALSPADINFEETLSTLRYADSAKRK 360  
 Db 291 -KKKKTDFIPYRDSVLTLMLRENGNSRTAMVAALSPADINYDETSLTRYADRAKQIK 349  
 Qy 361 NHAVYVEDPNAMIRKLELAQLRSKLGSGGGGGAGGSGGPVEESYPPDTPLEKQ-- 418  
 Db 350 CNAYVEDPNKLVRELKEEYTRKDLRLAQLGDLIDTNGSLT--SSPSCSLSSQW 407  
 Qy 419 ---IVSIQOPDATVYKKSKAEIVDELQNGSEKLYRLDNLQTWEEKLAKTEIHKREAALEE 475  
 Db 408 LTVSTIQ--ERIMSTPGCEAREIKSEKIIAELNETWEEKLAKTEIARKEREALAE 465  
 Qy 476 LGISIEK--GPGVPSHKEMPHLVNSDDPLLAECVLVYIKPGQTRVGNVNDQATRL 533  
 Db 466 MGVAIREDGTLGVSPKPHLVNLMDPLMSCECLLYIKGITRVGQADNERQIVL 525  
 Qy 534 NGSKILKEHCTFEV-----DNVTVIPVKEAAMVWYNDIKPTRLRSYRIILGDFHI 588  
 Db 526 SGAHKEEHCIIPERSNSGSEVITLPECSERSETYNGKVSQPVQLRSGNRIIMGNHV 585  
 Qy 589 FRNHPPEARACR-----OQSLNHSVTSNQ-----LGSP 619  
 Db 586 FRNHPQARHREKTPSAETSPVDMTFQRELLEKQIDMKQEKRLQDNELIYKK 645  
 Qy 620 APCRHTLSKAGSDAGDSRDS----- 643  
 Db 646 EKEEADLLQORLDADSDGSDSKSCSESMKLTSLREKLPSPKLTIVYKKGPISS 705  
 Qy 644 -----PLPHFR--GKDSYAFREAAAILGDKITSLTD-----DELDALFD 686  
 Db 705 GKREPIKMYQIPORRLSKDSKWTTISDLQIAQVKEICYEVA-LNDRFHSQETEAAL 764

Qy 687 DVQKARVVRCLVEDNEDSDSSSFVPRDKYKSNGTINDFSLDTALTPGTPRSD-- 742  
 Db 765 VKMELCAMYCKKIDPNE-RDSHRAY-ADWDVTGVGDEKIEDYMATGKGSTVDVLLK 822  
 Qy 743 -DGDALFPDCKSKSDASVNDVEELRQQAQMEAL 777  
 Db 823 IDKLEDILQEVYKQNNKMDKEIKVLRNKLMEKVL 858  
 RESULT 7  
 US-09-718-842-2  
 ; Sequence 2, Application US/09718842  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; TITLE OF INVENTION: Novel motor proteins and methods for  
 ; FILE REFERENCE: 1021  
 ; CURRENT APPLICATION NUMBER: US/09/718 842  
 ; CURRENT FILING DATE: 2000-11-22  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1153  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-718-842-2

Query Match 42.0%; Score 1691; DB 21; Length 1153;  
 Best Local Similarity 43.6%; Pred. No. 9,6e-124;  
 Matches 382; Conservative 137; Mismatches 235; Indels 122; Gaps 19;

Qy 4 GGNIKVVRVPFNAREIDRCAKIVRMENQOTILPPPGAEEKARKSKGTINDGPKAFA 63  
 Db 3 GASVAVVVRPFPNSBETSKESKCIIOQNGNSTSIINPNKE-----APKFS 51  
 Qy 64 FDRSYNSF-DKNAPYARQEDFDLGVPLDNAPKNCIFAYGOTGSKSYSMGTYG 122  
 Db 52 FDSYNSHTSPEDPCFASQNRVYNDIGKMLLHAFEGYVCIFAYGOTGAGKSYTMGQ 111  
 Qy 123 KEH--GVIPRICQDMFRINELQDKMLTCVVEVSYLETYNERYVDLLNPSTKGNLKVRE 180  
 Db 112 EESQAGIIPOLCELEFKIND-NCNEEMSYSEVSYMEYTCERYVDLLNPKNGLRVRE 170  
 Qy 181 HPSTGPIYVEDLAKLVRSFQENLDGNCARIVTAATNMNTHSSRHAVFTLTQKWH 240  
 Db 171 HPLGPIYVEDLSKLVSTYDIALMDAGNKARTVAATNMNTHSSRHAVFTIVTQKH 230  
 Qy 241 DEETKMDTKVAKISLVDLAGESEATSGATGAKEGAEINRSLSLTGLRVIAALADMS 300  
 Db 231 DNETNLSTKVSISLVDLAGESEATSGATGAKEGAEINRSLSLTGLKVISALAEVSK 290  
 Qy 301 GKOKNOLVPRDSVLTLMLKDSLGGSNTAMIAALSPADINFEETLSTLRYADSAKRK 360  
 Db 291 -KKKKTDFIPYRDSVLTLMLRENGNSRTAMVAALSPADINYDETSLTRYADRAKQIK 349  
 Qy 361 NHAVYVEDPNAMIRKLELAQLRSKLGSGGGGGAGGSGGPVEESYPPDTPLEKQ-- 418  
 Db 350 CNAYVEDPNKLVRELKEEYTRKDLRLAQLGDLIDTNGSLT--SSPSCSLSSQW 407  
 Qy 419 ---IVSIQOPDATVYKKSKAEIVDELQNGSEKLYRLDNLQTWEEKLAKTEIHKREAALEE 475  
 Db 408 LTVSTIQ--ERIMSTPGCEAREIKSEKIIAELNETWEEKLAKTEIARKEREALAE 465  
 Qy 476 LGISIEK--GPGVPSHKEMPHLVNSDDPLLAECVLVYIKPGQTRVGNVNDQATRL 533  
 Db 466 MGVAIREDGTLGVSPKPHLVNLMDPLMSCECLLYIKGITRVGQADNERQIVL 525  
 Qy 534 NGSKILKEHCTFEV-----DNVTVIPVKEAAMVWYNDIKPTRLRSYRIILGDFHI 588  
 Db 526 SGAHKEEHCIIPERSNSGSEVITLPECSERSETYNGKVSQPVQLRSGNRIIMGNHV 585

```

QY 589 FRNHPEARAER- - - - -OEOSLRHSHVTSNQ- - - - -LGSP 619
DB 586 FRNHPEARAERKTPSAETPSPVDWTFQAORELLEKQCIDMKOEMKRLQEMILYKK 645
QY 620 APCRHDRTLSKAGSDADGSRSD- - - - - 643
DB 646 EKEEADLLLEQRLDADSDGSDSDKRSCEBSWKLTLSLREKLPPSKLQTVKKCGLPSS 705
QY 644 - - - - -PLPHFR- - - - -GKSDWFYARREASATILGDKQISHLTD- - - - -DELDALFD 686
DB 706 CKKREPKWQIPORRLSKDSKWTYISDLKTOAVEICEVEA- - - - -LNDFRHSRQIEALAI 764
QY 687 DVOKARAVRGLVEDNEDSDSSFPFVRDYKYSNGTIDNFSLDTAITMPTFRSD- - - - -742
DB 765 VKMKELCANYGKKDPNE- - - - -ROSNRAV- - - - -ARDVWOTVGVGDEKIEDVNATGKGSTVDYDLKVH 822
QY 743 - - - - -DGDALFFGDKKQSDASNDVDELROQOAMEAL 777
DB 823 IDKLEDILQYKVKQNNKMADEIKVLRLNKLKMKVYL 858

RESULT 8
US-09-718-563-6
; Sequence 6, Application US/09718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-6

```

```

Query Match 41.7%; Score 1680.5; DB 21; Length 1770;
Best Local Similarity 47.3%; Pred. No. 1.3e-122;
Matches 353; Conservative 124; Mismatches 175; Indels 95; Gaps 14;

```

```

QY 4 GGNKVVVRPFPNAREIDRGAKCIVRMGNOQTILTPPGABEARKSGKTMIDGPKAFA 63
DB 3 GASKVAVRVPFPNSRSTKESKICIQOQNSTSIINPKPKE- - - - -APKPS 51
QY 64 FDSYWSF- - - - -DKNAPNVAQEDLPDGLVPLLDNAFKYNNCIPAYGOTSGKSYSMGVG 122
DB 52 FDSYWSHTSPEDPCFASQNRVTNDIGREMLHAFEGINVCIPAYGOTGAGKSYTMGMQ 111
QY 123 KEH- - - - -GVIPRICODMFRINELOKKNLTCTVEYSYLEIYNRVEDLLNPSTKGNLYRE 180
DB 112 ESQAGIIPQUCLELFKIND- - - - -NCNEENSYSEVSYMEIYCEYRVDLLAPKKNGLKYRE 170
QY 181 HPSTGYPVEDLAKLVRSFOPIENLMDGKAKTVAATNNETSSRHAVFTLTOKWH 240
DB 171 HPLGYPVEDLSKAVTSITDIDLNDGKNKRTVAATNNETSSRHAVFTLTOKWH 230
QY 241 DEETKMDTEKVKISIVDLGASERATSGATGARKKEGATINRSLTGLRVIALADMS 300
DB 231 DNETNLSTKYSKISIVDLGASERADSGATGARKKEGANIKSLITLGVLSALAEVK 290
QY 301 GKOKKQNLVPRVSVLTWLLKDSLGNSMTAATAISPADINFEETLSLTVADSAKRIK 360
DB 291 - - - - -KKKTDIFIPYRSVLTWLLRENLSGNSRTAMVAALSPADINFEETLSLTVADSAKRIK 349
QY 361 NHAVNEDPNARMRELKEELAQRLSKSSGGGGGAGCGSPVPSYPTDPLEK- - - - -418
DB 350 CNAVINEPNAKLVRELKEEVTRLDLRAQGLGDIIDTSMGSLT- - - - -SSPSSCSLSQVG 407

```

```

QY 419 - - - - -IVSIQDPDVTAKMSKAEIVYEOLNQSEKLYRDLNQTWEKLAQTEEHKEREAALE 475
DB 408 LTVTSIQ- - - - -ERIMSTPGEEATERKESEK IIAELNETWEKLRKTEATRMEREALLAE 465
QY 476 LGTISIEK- - - - -GPVGPYHSKEMPHLVNLSDDPLAECGLVYNNIKPGQTRGVNNOYTOAEIRL 533
DB 466 NGVAIREDDGGTLGVFSPKKTPLVNLNEDPLMSECLLYIKDGITRVGQADARRQDIYL 525
QY 534 NGSKILKEHCTFENV- - - - -DNVYTVIPNSKAAVMVNRIDKPTPLRSYRILGLDFHI 588
DB 526 SGAHKEEBHICIFRSRNSGSEVITVLEPCERSETVYVNGKVSOPVQLRSNGRIIMGNHV 585
QY 589 FRNHPEARAERQEOQLRUHSYTSNQLGSPAGPGRHDTLSKAGSDADGSDSSSPLPHF 648
DB 586 FRNHPEARAERK- - - - - 606
QY 649 RGKSDWFYARREASATILGDKQ- - - - -ISHLTDDELALFD- - - - -687
DB 607 PSEPVWTFQAORELLEK- - - - -OGIDMKOEMKRLQEMILYKKEEADLLLEQRLDYESKL 665
QY 688 - - - - -VOKARAVRGLVEDNEDSDSSFPF 712
DB 666 QALQKQVETRSLAAETTEEEEEEVPP 692

RESULT 9
US-09-718-804-6
; Sequence 6, Application US/09718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-6

```

```

Query Match 41.7%; Score 1680.5; DB 21; Length 1770;
Best Local Similarity 47.3%; Pred. No. 1.3e-122;
Matches 353; Conservative 124; Mismatches 175; Indels 95; Gaps 14;

```

```

QY 4 GGNKVVVRPFPNAREIDRGAKCIVRMGNOQTILTPPGABEARKSGKTMIDGPKAFA 63
DB 3 GASKVAVRVPFPNSRSTKESKICIQOQNSTSIINPKPKE- - - - -APKPS 51
QY 64 FDSYWSF- - - - -DKNAPNVAQEDLPDGLVPLLDNAFKYNNCIPAYGOTSGKSYSMGVG 122
DB 52 FDSYWSHTSPEDPCFASQNRVTNDIGREMLHAFEGINVCIPAYGOTGAGKSYTMGMQ 111
QY 123 KEH- - - - -GVIPRICODMFRINELOKKNLTCTVEYSYLEIYNRVEDLLNPSTKGNLYRE 180
DB 112 ESQAGIIPQUCLELFKIND- - - - -NCNEENSYSEVSYMEIYCEYRVDLLAPKKNGLKYRE 170
QY 181 HPSTGYPVEDLAKLVRSFOPIENLMDGKAKTVAATNNETSSRHAVFTLTOKWH 240
DB 171 HPLGYPVEDLSKAVTSITDIDLNDGKNKRTVAATNNETSSRHAVFTLTOKWH 230
QY 241 DEETKMDTEKVKISIVDLGASERATSGATGARKKEGATINRSLTGLRVIALADMS 300
DB 231 DNETNLSTKYSKISIVDLGASERADSGATGARKKEGANIKSLITLGVLSALAEVK 290
QY 301 GKOKKQNLVPRVSVLTWLLKDSLGNSMTAATAISPADINFEETLSLTVADSAKRIK 360
DB 291 - - - - -KKKTDIFIPYRSVLTWLLRENLSGNSRTAMVAALSPADINFEETLSLTVADSAKRIK 349

```

```

Qy 361 NHAVNEDPNARMIRELKEALQRLSKLQSSGGGGAGGSGGPPVESTPDPPLKEQ-- 418
Db 350 CNVYNEDPNMAKIVRELKEEYTRKLDLLRAAGLDIDTSMGSLT--SSPSSCULSSQVG 407
Qy 419 ---IVSYQDPATYKMSAEIVQDLNOSKLYRDLNQTWPEKLAETETHEKREAALEE 475
Db 408 LTVSYTIO--ERIMSTPGGEATERLKESEKIAELNETWEKLRKTAERMRERALLAE 465
Qy 476 LGISTEK--GFVGPVHSKMPHLVNSDDPLAECVLVYNIKPGQTRVGNVNDQTQAEIRL 533
Db 466 MGVAIREDGGTLGVSPKTPHLVNLNEDPLMSECLLYIKDGITRVGQADAEERQDVL 525
Qy 534 NSGKILKEHCTFENY-----DNVVTIVPEKAAMVNVGRIDKPTRLRSGYRIILGDPHI 588
Db 526 SGAHKEHCICFRSERSNGEVITVLPCESETYVNGKRVSQVQLRSGNRIIMGNHNV 585
Qy 589 FRFNHPPEARARQESQLRHSVYNSQLGSPAPGRHDRTLKSGADGADGSDRSPLPHF 648
Db 586 FRFNHPQQAAREK-----ISHLTDDDELALFDD-----TFSAET 606
Qy 649 RGKDSWYFAREASAILGLDQK-----ISHLTDDDELALFDD----- 687
Db 607 PSEPVDMTFAQRELLEK--QGIDMKQMEKRLQEMELIYKKEKEADLLLEQRLDYESKL 665
Qy 688 --VOKARAVRGCLVEDNEDSDQSFP 712
Db 666 QALQXQVETRSLAAETTEEBEEBEEVP 692

RESULT 10
US-09-718-842-6
; Sequence 5, Application US/09718842
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; their use
; REFERENCE: 02/010000
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-6

```

```

Query Match 41.7%; Score 1680.5; DB 21; Length 1770;
Best Local Similarity 47.3%; Pred. No. 1.3e-122;
Matches 353; Conservative 124; Mismatches 175; Indels 95; Gaps 14;

```

```

Qy 4 GGNIKVVRVFPNAREIDRGAKCIYVWEGNQTILTPPGAEEKARKSGKTIHOGPKAFA 63
Db 3 GASKVAVRVFPNAREIDRGAKCIYVWEGNQTILTPPGAEEKARKSGKTIHOGPKAFA 63
Qy 64 FDSYKASF--DKMANVAREDELFDGLVPLLDNAKGYNVCIFAYGOTGSGKSYSMWVG 122
Db 52 FDSYKASHSPDCAFQSNRYVNDICKEMLLHAFEGYVVCIFAYGOTGAGKSYTMAGQ 111
Qy 123 KEH--GVIPICQDMFRRIEQLQDKNLTCTVEVSYLLETYNERVROLLNPSTKGNLKV 180
Db 112 BESQAGIIPOLCEELFEKIND--NCNEMSYSVEVSYMEIYCERVROLLNPKNKGNLKV 170
Qy 181 HPSTQPVVEDLAKLVRSFOETENIDECNKARTVAATNMNETSSRSHAVFTLTQKWH 240
Db 171 HPLGPGVEDLSKLVSYTDIADLDAGNKARTVAATNMNETSSRSHAVFTIVPTQKH 230
Qy 241 DEETKMDTEKAKTISLVLAGSERATSGATGKRLKEGAEINRSLSTLGRVIAALDMS 300
Db 231 DNETNLSTKYSKISLVLAGSERADSTGKTRLKEGANKINKSLTLTGKVISALAEVSK 290

```

```

Qy 301 GKQKNQLVPYRDSVLTMLLKSDLSGNSMTAMIAAISPADINFEETLSTLRVADSARIK 360
Db 291 -KKKKTDIFPYRDSVLTMLLNGNSFRNVAALSPADINTDETSLTRYADRAKQIK 349
Qy 361 NHAVNEDPNARMIRELKEALQRLSKLQSSGGGGAGGSGGPPVESTPDPPLKEQ-- 418
Db 350 CNVYNEDPNMAKIVRELKEEYTRKLDLLRAAGLDIDTSMGSLT--SSPSSCULSSQVG 407
Qy 419 ---IVSYQDPATYKMSAEIVQDLNOSKLYRDLNQTWPEKLAETETHEKREAALEE 475
Db 408 LTVSYTIO--ERIMSTPGGEATERLKESEKIAELNETWEKLRKTAERMRERALLAE 465
Qy 476 LGISTEK--GFVGPVHSKMPHLVNSDDPLAECVLVYNIKPGQTRVGNVNDQTQAEIRL 533
Db 466 MGVAIREDGGTLGVSPKTPHLVNLNEDPLMSECLLYIKDGITRVGQADAEERQDVL 525
Qy 534 NSGKILKEHCTFENY-----DNVVTIVPEKAAMVNVGRIDKPTRLRSGYRIILGDPHI 588
Db 526 SGAHKEHCICFRSERSNGEVITVLPCESETYVNGKRVSQVQLRSGNRIIMGNHNV 585
Qy 589 FRFNHPPEARARQESQLRHSVYNSQLGSPAPGRHDRTLKSGADGADGSDRSPLPHF 648
Db 586 FRFNHPQQAAREK-----ISHLTDDDELALFDD-----TFSAET 606
Qy 649 RGKDSWYFAREASAILGLDQK-----ISHLTDDDELALFDD----- 687
Db 607 PSEPVDMTFAQRELLEK--QGIDMKQMEKRLQEMELIYKKEKEADLLLEQRLDYESKL 665
Qy 688 --VOKARAVRGCLVEDNEDSDQSFP 712
Db 666 QALQXQVETRSLAAETTEEBEEBEEVP 692

RESULT 11
US-60-258-275-444
; Sequence 444, Application US/60258275
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS.
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001026-PROV
; CURRENT APPLICATION NUMBER: US/60/258, 275
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 717
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 444
; LENGTH: 1793
; TYPE: PRT
; ORGANISM: HUMAN
US-60-258-275-444

```

```

Query Match 41.4%; Score 1669.5; DB 23; Length 1793;
Best Local Similarity 46.8%; Pred. No. 9.8e-122;
Matches 352; Conservative 125; Mismatches 166; Indels 109; Gaps 15;

```

```

Qy 4 GGNIKVVRVFPNAREIDRGAKCIYVWEGNQTILTPPGAEEKARKSGKTIHOGPKAFA 63
Db 9 GASKVAVRVFPNAREIDRGAKCIYVWEGNQTILTPPGAEEKARKSGKTIHOGPKAFA 63
Qy 64 FDSYKASF--DKMANVAREDELFDGLVPLLDNAKGYNVCIFAYGOTGSGKSYSMWVG 120
Db 58 FDSYKASHSPDCAFQSNRYVNDICKEMLLHAFEGYVVCIFAYGOTGAGKSYTMAG 115
Qy 121 YCK--EHGVIPRICQDMFRRIEQLQDKNLTCTVEVSYLLETYNERVROLLNPSTKGNLKV 178
Db 116 KQEKDOOGIIPOLCEELFSRINDTND--NMYSYSEVSYMEIYCERVROLLNPKNKGNLKV 174
Qy 179 REHSTGPGVEDLAKLVRSFOETENIDECNKARTVAATNMNETSSRSHAVFTLTQK 238
Db 175 REHPLGPGVEDLSKLVSYTDIADLDAGNKARTVAATNMNETSSRSHAVFTIQTQK 234

```

QY	239	WHDDETAKOTKRYKAKISLVLDLAGSERATSTGATGAKLKEGAEINRSLSTLGRVIAALADM	298
DB	239		
DB	235	RHDAETNITTKRYKISLVLDLAGSERADSTGAKTRKEGANINRSLSTLGRVIAALADM	294
DB	235		
QY	299	SSG-----KQKNOLVPYRDSVLTKLSDLGSGNTAMTATSPADINPEETLSTLYA	353
DB	299		
DB	295	DGSPHNKAKKRTDFPIYRDSVLTLURENGLGNSRTAMVALSPADINDETSLSTLYA	354
DB	295		
QY	354	DSAREINNAVNPEDNPAWMTRELUKELAQLRSLKLOSSGGGGGSGGSGGVPEVPPDT	413
DB	354		
DB	355	DRAKQICRNAVINEPNNKLRELKQDEVYRLDLTLTAAGLG-----DIT	398
DB	355		
QY	414	PLEKQIVSTQOEDPATVKKMS-----KAEIVPOLNOSKLYRDLNOTW	455
DB	414		
DB	399	DMTNALVGM-SPSSLSLSSRAASVSSLHERILLFAPGSEERLKEKTEKIIAELNETW	457
DB	399		
QY	456	EKLAKTEETHKREAALEELGISLEK--GFVGPYHSKEMPHLVNLSODPLIARCIIVNI	513
DB	456		
DB	458	EKLKRTAEATMRERALLAEMGVAMREDGGTLGVFPSPKTKPLVLNLEDPLMSECLLYI	517
DB	458		
QY	514	KFGOTRVGVNQNODTAERLNGSKILKEHCTFEN-----VDNVVTVIPNKAAVNGVR	568
DB	514		
DB	518	KDQITRVGREDDERRQDVLVLSGHFIEKHCVFSDRGSGSEAVVTLFECGADTVNGKK	577
DB	518		
QY	569	IDKPTLRSGYRITLGDHFIFRPNHPEAPAREOESLRLHVSNTSOLGSPAPGRHRTL	628
DB	569		
DB	578	VTEPSILRSGNIITGKSHVFRPNHPEAOERER-----	612
DB	578		
QY	629	STAGSDADGSDRSODPLHFRCGKSDWVFYARREASATILGDKQKLSHLTDDLDALFDV	688
DB	629		
DB	613	-----TPCAETPAEVDMAFAPARELLK-OGIDMK--QEMBOIRLOELEDQY	655
DB	613		
QY	689	QKARAVRGLVEDNEDSDSQSFPPYDKTMYN	720
DB	689		
DB	656	RREREATYILLE-QORLDYESKLEALQKQMD	686
DB	656		
RESULT 12			
US-09-718-563-8			
: Sequence 8, Application US/09718563			
: GENERAL INFORMATION:			
: APPLICANT: Peraud, Christophe			
: TITLE OF INVENTION: Novel motor proteins and methods for			
: TITLE OF INVENTION: their use			
: FILE REFERENCE: 1021			
: CURRENT APPLICATION NUMBER: US/09/718,563			
: CURRENT FILING DATE: 2000-11-22			
: NUMBER OF SEQ ID NOS: 10			
: SOFTWARE: FastSeq for Windows Version 4.0			
: SEQ ID NO 8			
: LENGTH: 1816			
: TYPE: PRT			
: ORGANISM: Human			
US-09-718-563-8			

Query Match 41.4%; Score 1668.5; DB 21; Length 1816;  
Best Local Similarity 45.0%; Pred. No. 1.2e-121;  
Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;

[illegible][illegible]

```

Query Match      41.4%; Score 1668.5; DB 21; Length 1816;
Best Local Similarity 45.0%; Pred. No. 1.2e-121;
Matches 336; Conservative 126; Mismatches 172; Indels 137; Gaps 16;

QY 4 GGNIKVYVRVPNAREIDRCACIVRMEGNQITLTPPPGAEAKRKSGKTIIDMGKPAFA 63
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 GASVAGAVYVRPNRETSKFSKCTIOMGNSTSTINPKNPKR-----ADKPS 51
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Query Match 41.4%; Score 1668.5; DB 21; Length 1816;  
Best Local Similarity 45.0%; Pred. No.1.e2-121;  
Matches 356; Conservative 126; Mismatches 172; Indels 137;

QY 4 GGNTVVVVVPPFNAREIDRGACIVCMQNETIUTPPGAEKARKSKTINDGKPAF  
db 3 GASVKVAVVPPNPSBESKSKSCTTCMGNTSTINPKNPF-----ADKSP

Qy 64 FDRSYWSE--DKNAPNYARQEDFQDGLVPLLDNAFKYNNCIFAYGQTGSKYSNMGYG 122  
Db 52 FDSYWSHSTSPEDPCFASQNRVYINDIOKEMLLHAFEGYVNCIFAYGQTGACKSYTMGKQ 111  
Qy 123 KEH--GVIPRIQDPMFRINELQOKNUTCTVEVSYLYNERNVRLDLPSTAGNLKAVRE 180  
Db 112 ESSQAGVITPOLCEBEFPEKIND--NCNEEMSYSVESVYELCYERVRLDLPKNKNLVR 170  
Qy 181 HPSTGPYVEDLAKLVYRSFOENIMDGNGKARTVAATNNETSSRSHAVFTLTQKHH 240  
Db 171 HPLGCPYVEDLSKLVNTYDIAOLMDAGNKARTVAATNNETSSRSHAVFTIVTQKHH 230  
Qy 241 DEETKMDTEVKAKISLVLDAGSERATSGATGARKEGAEINRSITLTGRVIAALADM-- 298  
Db 231 DNETNLSTEVKSKI SLVDLAGSERADSTGATGARKEGANINKSLTTLGKVISALAEVDN 290  
Qy 299 ---SSGOKKXNOLVYPYRDSVLTWLLKSLGNSMTAMIAALSPADINFEETLSTLYADS 355  
Db 291 CTSKSKKKKKTKDPIPYRDSVLTWLLRNLGNSRTAMVAALSPADINDYDTLSTLYADR 350  
Qy 356 AKRKHNHVNEDPNMARIRELKEBELAOLRSKLOSSGGG-----GGAG----- 399  
Db 351 AKQIKCNVAVINEDPNMARELKEBELAOLRAQGLDIDIDPLDIDYSGSGSKYLK 410  
Qy 400 -----GGSGPVESYPPDTPLEKQ-----IVTSIQPDATYKK 431  
Db 411 DFQNNKHRYLLASENQRPFGHSTAGMSLTSS--PSSCSLSQVGLTSTVSIQ--ERINST 467  
Qy 432 MSKAEIVQLOLNOSEKLYRLDQNTWEKLAETIEHKREAALEGLISIEK--GFVGPVH 489  
Db 468 PGGEAEERLKESEKIITAEELNETWEKLRKTEAIRMEREALLAEMGVAIREDGTLGVES 527  
Qy 490 SKEMPHLVNLSDDPLAECVLVYIKPQTRVGNVQDQAEIRLNGSKILKEHCTFENV- 548  
Db 528 PKTTPHVLNEDPLMSECLLYIKDGIITRVGQADERRQDVLGSAHIKEBICFRSER 587  
Qy 549 ----DNVYTVPNKAAVWVNRIDKPTLRFSGYRIILGDPHIFRNHPPEARAEQBO 604  
Db 588 SNSGEVITLPECESETTYVNGRVQSPQVLRSGNRITMGKNHVRFNHPPEARAEK- 646  
Qy 605 SLLRHVYTNLSQSPAGPHDRITLSKAGSDAGDSRSDPLPHFGKDSDFMYARREAS 664  
Db 647 -----TPSAETPSEPVDMTFAOELLE 668  
Qy 665 AILGLDOK-----ISHLTDDELDALEFD-----VOKARAVRGLVED 701  
Db 669 K-QGIDMKQMEKRLQEMETLYKKEKEEADILLEQORLDYESKLOALQKQVETRSAAET 727  
Qy 702 NEDSDSSQSPF 712  
Db 728 TEEEEEVEVP 738

RESULT 14  
US-09-718-842-8  
; Sequence 8, Application US/09718842  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; FILE REFERENCE: Their use  
; CURRENT APPLICATION NUMBER: US/09/718,842  
; CURRENT FILING DATE: 2000-11-22  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 1816  
; TYPE: PRT  
; ORGANISM: Human  
US-09-718-842-8

Query Match 41.4%; Score 1668.5; DB 21; Length 1816;  
Best Local Similarity 45.0%; Pred. No. 1,2e-121;  
Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;  
Qy 4 GNRIKVVYVRPFPNAREIDRGAKCIVRMGEQNTILTPPGAEBEAKRKSGKTMIDGPKAFA 63  
Db 3 GASVYAVAVYRFPNAREIDRGAKCIVRMGEQNTILTPPGAEBEAKRKSGKTMIDGPKAFA 63  
Qy 64 FDRSYWSE--DKNAPNYARQEDFQDGLVPLLDNAFKYNNCIFAYGQTGSKYSNMGYG 122  
Db 52 FDSYWSHSTSPEDPCFASQNRVYINDIOKEMLLHAFEGYVNCIFAYGQTGACKSYTMGKQ 111  
Qy 123 KEH--GVIPRIQDPMFRINELQOKNUTCTVEVSYLYNERNVRLDLPSTAGNLKAVRE 180  
Db 112 ESSQAGVITPOLCEBEFPEKIND--NCNEEMSYSVESVYELCYERVRLDLPKNKNLVR 170  
Qy 181 HPSTGPYVEDLAKLVYRSFOENIMDGNGKARTVAATNNETSSRSHAVFTLTQKHH 240  
Db 171 HPLGCPYVEDLSKLVNTYDIAOLMDAGNKARTVAATNNETSSRSHAVFTIVTQKHH 230  
Qy 241 DEETKMDTEVKAKISLVLDAGSERATSGATGARKEGAEINRSITLTGRVIAALADM-- 298  
Db 231 DNETNLSTEVKSKI SLVDLAGSERADSTGATGARKEGANINKSLTTLGKVISALAEVDN 290  
Qy 299 ---SSGOKKXNOLVYPYRDSVLTWLLKSLGNSMTAMIAALSPADINFEETLSTLYADS 355  
Db 291 CTSKSKKKKKTKDPIPYRDSVLTWLLRNLGNSRTAMVAALSPADINDYDTLSTLYADR 350  
Qy 356 AKRKHNHVNEDPNMARIRELKEBELAOLRSKLOSSGGG-----GGAG----- 399  
Db 351 AKQIKCNVAVINEDPNMARELKEBELAOLRAQGLDIDIDPLDIDYSGSGSKYLK 410  
Qy 400 -----GGSGPVESYPPDTPLEKQ-----IVTSIQPDATYKK 431  
Db 411 DFQNNKHRYLLASENQRPFGHSTAGMSLTSS--PSSCSLSQVGLTSTVSIQ--ERINST 467  
Qy 432 MSKAEIVQLOLNOSEKLYRLDQNTWEKLAETIEHKREAALEGLISIEK--GFVGPVH 489  
Db 468 PGGEAEERLKESEKIITAEELNETWEKLRKTEAIRMEREALLAEMGVAIREDGTLGVES 527  
Qy 490 SKEMPHLVNLSDDPLAECVLVYIKPQTRVGNVQDQAEIRLNGSKILKEHCTFENV- 548  
Db 528 PKTTPHVLNEDPLMSECLLYIKDGIITRVGQADERRQDVLGSAHIKEBICFRSER 587  
Qy 549 ----DNVYTVPNKAAVWVNRIDKPTLRFSGYRIILGDPHIFRNHPPEARAEQBO 604  
Db 588 SNSGEVITLPECESETTYVNGRVQSPQVLRSGNRITMGKNHVRFNHPPEARAEK- 646  
Qy 605 SLLRHVYTNLSQSPAGPHDRITLSKAGSDAGDSRSDPLPHFGKDSDFMYARREAS 664  
Db 647 -----TPSAETPSEPVDMTFAOELLE 668  
Qy 665 AILGLDOK-----ISHLTDDELDALEFD-----VOKARAVRGLVED 701  
Db 669 K-QGIDMKQMEKRLQEMETLYKKEKEEADILLEQORLDYESKLOALQKQVETRSAAET 727  
Qy 702 NEDSDSSQSPF 712  
Db 728 TEEEEEVEVP 738  
RESULT 15  
US-60130217-6  
; Sequence 16, Application US/60130217  
; GENERAL INFORMATION:  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Dufauere-Gare, Isabelle  
; APPLICANT: Grel, Pascale  
; TITLE OF INVENTION: DNA encoding a novel kinesin-like protein  
; FILE REFERENCE: GENSET.060PR  
; CURRENT APPLICATION NUMBER: US/60/130,217  
; CURRENT FILING DATE: 1999-04-20

```

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent.pm
; SEQ ID NO 6
; LENGTH: 1816
; TYPE: PR1
; ORGANISM: homo sapiens
US-60-130-217-6

```

Query Match	41 48;	Score	1668, 5;	DB	23;	Length	1816;
Best Local Similarity	45.08;	Prod. No.	1.2e-172;				
Matches	356;	Conservative	126;	Mismatches	171;	Indels	137;
Gaps	165;						
QY	4	GGNIKVYVRPENAREIDRGACACTVMEGNOTILTPPGAEKARKSKGTIMDQPKAFA	63				
DB	3	GASVKVAVVRPENSRETSKESKICIIQGNSTSIINPKPE	51				
QY	64	FRDSYWSF-DKNAPYAROEIDFDGLGVPLLDNAFGKYNKNCIFAYCGTGSGKSYVMCQY	122				
DB	52	FDYSYWSHTSPEDPCFASONRVYNDIGKEMLLHAFEGYVNCIFAYCGTAGCKSYTMCKQ	111				
QY	123	KEH-GVIPRICQDMFRINELQKKNITCTVEVSYLEINERVRLDNLNSTGNLKVRE	180				
DB	112	EESQAGIIPOLCELPFXIND-NCNEEMSYSVEVSMEICVRVRLDNLNPKNGNLVRE	170				
QY	181	HPSTGPYVEDLAKLVVRSFOEINLEMDGNKARTVAATNNNETSSRSHAVETLITQKWH	240				
DB	171	HPLLGPYVEDLSLAVTSYTDIADLMDAGNARTVAATNNNETSSRSHAVETVTFQKH	230				
QY	241	DEETKMDTEKVARISLVDLAGSERATSYGATGARLKEGAEINRSASTLGRVIAALDM-	298				
DB	231	DNETNLSTKVKYSISLVDLAGSERADSTGARGRLKEGANINSLTILGKYVISALAEVDN	290				
QY	299	---SSGOKKNQLVYRDSVLTWLLKDSGNSMTAIAISPADINFEETLSTLYADS	355				
DB	291	CTSCKKKKKTTDPIYRDSVLTWLLRNLGNSRTAMTAAISPADINDETSLTRYADR	350				
QY	356	AKRIKHNAVNEPDNARMRELKEELQASRLQSGGGG	359				
DB	351	AQIKCNAVINEDPNAKLVRELKEETLKLDRAGGLGDIIDPLDDYSGSGSKYLK	410				
QY	400	---GSGPVBEESYPPDTPLEKQ----	431				
DB	411	DFONKHKRYLASENORGHFSTASMSLTSYSS-PSCSLSQVGLTSVTSIQ--RRIMST	467				
QY	432	MSKAEIVQNLQSEKLYRDLNQTWTEKIATBIEHKREALBELGISIEK--GFVGPYH	489				
DB	468	PGGEATIERKSEKITAELNETWEELURKYEATIRMEAREALBAEMGVAIREDGGTLGVFS	527				
QY	490	KXEMHVLNLSDDPLACLVLNYPGQTVRVGNVDQTAIRLNGSKILKEHCTFNW-	548				
DB	528	PXTFPLYNEDPLMSCEVYIKDGTPELRSYRIITLGDPIHFPHPEARAEARQDQ	604				
QY	549	---DNVYTVNEKAAWNGVRIKDTPELRSYRIITLGDPIHFPHPEARAEARQDQ	604				
DB	588	SNSGEVIVTLEPCERSERTYNGKRVSQPVOLRSNRIIMKNGNVPFPHPEARAEARCK-	646				
QY	605	SLLRHSVTNSQLGSPACGRHRTLSKAGSDADCSRSDPLPHFRGKSDWFVAREAS	664				
DB	647	---TPSAETPSEVDMTAQORELLE	668				
QY	665	AULGLDOK-----ISLHTDDELDALFDO-----	701				
DB	669	K-GQDMKQWEKRLQEMEILYKKEKEADLLLEQQLDYESKLQALQVETRSIAAET	727				
QY	702	NEDSDSQSSFP	712				
DB	728	TEEEEEEVFP	738				



**THIS PAGE BLANK (USPTO)**



GenCores version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:11:29 : Search time 7.94 Seconds  
(without alignments)  
311.456 Million cell updates/sec

Title: US-09-235-416-1  
Perfect score: 4030  
Sequence: 1 MSGGNKKVVRPFNARE.....ELRQQAOMEALKTAKQEF 784  
Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5  
Searched: 17899 seqs, 3154390 residues

Total number of hits satisfying chosen parameters: 17899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_5/p/odata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_5/p/odata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_5/p/odata/1/paa/US01\_NEW\_COMB.pep.\*  
4: /cgn2\_5/p/odata/1/paa/US01\_NEW\_COMB.pep.\*  
5: /cgn2\_5/p/odata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_5/p/odata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673.5	16.7	513	5	US-09-724-519-6
2	672.5	16.7	575	5	US-09-724-519-6
3	672.5	16.7	575	5	US-09-724-519-6
4	628	15.6	358	5	US-09-724-519-6
5	618.5	15.3	382	5	US-09-724-519-6
6	613.5	12.7	790	5	US-09-723-153-2
7	457.5	11.4	338	5	US-09-723-153-2
8	133.5	3.3	1312	1	PCT-US01-08117-95
9	133.5	3.3	2816	1	PCT-US01-10484-145
10	130	3.2	992	6	US-60-248-505-1112
11	126.5	3.1	3078	6	US-60-248-505-895
12	122.5	3.0	619	5	US-09-739-449-10081
13	121.5	3.0	1108	5	US-09-739-449-9117
14	120	3.0	1833	4	US-08-945-567C-4
15	120	3.0	1992	4	US-08-945-567C-3
16	117	2.9	1179	5	US-09-792-024-89
17	115.5	2.9	1058	6	US-60-248-505-1230
18	114.5	2.8	1494	5	US-09-421-124-186
19	114.5	2.8	1494	5	US-09-421-124-186
20	113.5	2.8	1594	5	US-09-421-124-186
21	113.5	2.8	1379	5	US-09-421-124-184
22	113.5	2.8	1879	5	US-09-421-124-184
23	113.5	2.8	2137	6	US-60-248-505-990
24	113	2.8	1156	5	US-60-248-505-990
25	113	2.8	1206	6	US-60-248-505-1161
26	112.5	2.8	3159	6	US-60-248-823-85
27	110.5	2.7	2031	5	US-09-739-449-10549

28	110.5	2.7	2329	5	US-09-421-124-16
29	110.5	2.7	2329	5	US-09-421-124-16
30	107.5	2.7	3418	5	US-09-421-124-44
31	107.5	2.7	3418	5	US-09-421-124-44
32	107	2.7	934	5	US-09-707-468A-11
33	107	2.7	1167	5	US-09-739-449-9204
34	106	2.6	902	6	US-60-248-505-973
35	103.5	2.6	520	5	US-09-739-449-8981
36	103.5	2.6	1416	1	PCT-US01-08117-4
37	102.5	2.5	1953	6	US-60-248-505-1061
38	102	2.5	594	5	US-09-739-449-8956
39	100.5	2.5	591	5	US-09-739-449-8956
40	100	2.5	1535	5	US-09-421-124-185
41	100	2.5	1535	5	US-09-421-124-185
42	100	2.5	3170	6	US-60-248-505-909
43	99.5	2.5	496	5	US-09-819-386-4
44	99.5	2.5	922	5	US-09-739-449-10685
45	99	2.5	371	6	US-60-248-505-1198

ALIGNMENTS

RESULT 1  
US-09-724-519-6  
; Sequence 6, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finer, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Mak, John  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Methods of screening for modulators of cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; FILE REFERENCE: 1014A  
; CURRENT APPLICATION NUMBER: US/09/724,519  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 09/592,037  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR APPLICATION NUMBER: 09/428,156  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-519-6

Query Match  
Matches 187; Conservative 85; Mismatches 171; Indels 77; Gaps 18;  
Best Local Similarity 36.0%; Pred No. 2.9e-39;

QY	4	GNRIKVVVRPFNAREIDRGAKCIYVRMGNOTILTPPGAEKAKSKGTIMDGPKAFA 63
DB	16	GKNIQVVRCPFNLAERKASAHISVECD-----PVRKEVSVRTGGLADKSSRKYTT 67
QY	64	FRDSYWSFKNAPNAREIDFQDLGVPLLDNAFKYNNCTIFAYGTQSGKYSNMG--- 120
DB	68	FDWVGASTK-----QIDVRSYVVCPLDDEVINGNCTIFAYGTQSGKFTMEGRS 120
QY	121	---YGEH---GVPIPCQDMFRINELQDKNLCTVEVSYLEIYNERVRLNPST 172
DB	121	PNSEYTWEDPLAGIIPTRUFEKLT---DNCTEFSVKVSLLEYNEELFDLNPSS 176
QY	173	---KGMKVREHP---STGYDVEDAKLVRSPOINLMOEGNKARTVAATNNWTSRSH 228
DB	177	DVSERLQMDPDRNKGVIIGLEIETVHNKDEYVLEKGNARTTANTLWYSSRSH 236
QY	229	AVPTLTITQKHDEETKMDTE---KVAKISLDVLGASERATSGATGARGKAEINSL 285



Db 345 TLTSTLEYAHRANKNLNKPVPNQKLTKKALKEYTTEBIEBKRLDAAAREKNG----- 396  
QY 405 V---ESY-----PDPPLKQIVSVQOPDPAVKKMSKAEIVQLNQSSEKLYRD-----LNO 453  
Db 397 VYISEENFVWSKGLTVQSEIVEL-----TEKIGAVE--BELNRTVALFRONKNELQO 448  
QY 454 TWEEKLANTTEHKEEAALEPGLISTEK 482  
Db 449 CKSDLQNKQTOEL-ETTKHQLETKQLQVK 476  
RESULT 4  
US-09-724-519-10  
; Sequence 10, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finer, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakwicz, Roman  
; APPLICANT: Sak, John  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; FILE REFERENCE: 1014A  
; CURRENT APPLICATION NUMBER: US/09/724.519  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 09/592,037  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR APPLICATION NUMBER: 09/428,156  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-519-10

Query Match 15.6% Score 628; DB 5; Length 368;  
Best Local Similarity 40.6% Pred. No. 2.3e-36;  
Matches 155; Conservative 57; Mismatches 122; Indels 48; Gaps 10;

QY 4 GGNKVVYRVFPFNAREIDRGAKCIVRMGNOTILTPPGAEEKARKSGKTIIMDGPKAFA 63  
Db 16 GKNIOVVYRCRPFNLAERKASAHISVECD-----PYRKEVSVRTGGLADKSSRTYIT 67  
QY 64 FDRYSWSPKNAPYARQEDLPQDLGVPLLDNAFKNYNNCTIFAYGOTGSGKYSWMG--- 120  
Db 68 FDMVFGASTR-----QIDVYRSVYCPILDEVIMYNGTCTIFAYGOTGKTFTMEGRS 120  
QY 121 -----YKEH-----GVPRICODMFRINELQKNKUTCYVEVSYLEIYNERVOLLNPST 172  
Db 121 PNEYTWEEEDPLAGIIPRTLHQIPEKLT-----DNGTEFSYKVSLEIYNEELFDLLNPSS 176  
QY 173 --KGNLVKREHP--STGPYVEDLAKLVRSFOETENLMDGNKARTVVAATNNWETSSRSH 228  
Db 177 DVSERLQMFDDPRNKGVIKGLSEYTHVKNDEVYGLKGAAKRTTAATLNNAYSSRSH 236  
QY 229 AVFTLTITQKHDEETKMDE--KVAKISLVDLAGSERATSTGATGARLKEGAEINRSL 285  
Db 237 SVFESVTI---HMKETTIDGELVYKGLNLDVLAGSENIGRSGAVDKRAREAGNINSL 292  
QY 286 STLGRVIAALADMSSGSKKKNQJVPYRDSVLTMLKDSLAGNSMTAMTAISPADINFE 345  
Db 293 LTLGRVITAVE-----RTPHPVYRESKRLTLQDSLGGRTRTSTIATISPASLNLEE 345  
QY 346 TLTSTLEYAHRANKNLNKPVPNQKLTKKALKEYTTEBIEBKRLDAAAREKNG----- 396  
Db 346 TLTSTLEYAHRANKNLNKPVPNQKLTKKALKEYTTEBIEBKRLDAAAREKNG----- 396

RESULT 5

US-09-724-519-4  
; Sequence 4, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finer, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakwicz, Roman  
; APPLICANT: Sak, John  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; FILE REFERENCE: 1014A  
; CURRENT APPLICATION NUMBER: US/09/724.519  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 09/592,037  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR APPLICATION NUMBER: 09/428,156  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-519-4

Query Match 15.3% Score 618.5; DB 5; Length 382;  
Best Local Similarity 39.4% Pred. No. 1.1e-35;  
Matches 157; Conservative 61; Mismatches 125; Indels 55; Gaps 12;

QY 4 GGNKVVYRVFPFNAREIDRGAKCIVRMGNOTILTPPGAEEKARKSGKTIIMDGPKAFA 63  
Db 16 GKNIOVVYRCRPFNLAERKASAHISVECD-----PYRKEVSVRTGGLADKSSRTYIT 67  
QY 64 FDRYSWSPKNAPYARQEDLPQDLGVPLLDNAFKNYNNCTIFAYGOTGSGKYSWMG--- 120  
Db 68 FDMVFGASTR-----QIDVYRSVYCPILDEVIMYNGTCTIFAYGOTGKTFTMEGRS 120  
QY 121 -----YKEH-----GVPRICODMFRINELQKNKUTCYVEVSYLEIYNERVOLLNPST 172  
Db 121 PNEYTWEEEDPLAGIIPRTLHQIPEKLT-----DNGTEFSYKVSLEIYNEELFDLLNPSS 176  
QY 173 --KGNLVKREHP--STGPYVEDLAKLVRSFOETENLMDGNKARTVVAATNNWETSSRSH 228  
Db 177 DVSERLQMFDDPRNKGVIKGLSEYTHVKNDEVYGLKGAAKRTTAATLNNAYSSRSH 236  
QY 229 AVFTLTITQKHDEETKMDE--KVAKISLVDLAGSERATSTGATGARLKEGAEINRSL 285  
Db 237 SVFESVTI---HMKETTIDGELVYKGLNLDVLAGSENIGRSGAVDKRAREAGNINSL 292  
QY 286 STLGRVIAALADMSSGSKKKNQJVPYRDSVLTMLKDSLAGNSMTAMTAISPADINFE 345  
Db 293 LTLGRVITAVE-----RTPHPVYRESKRLTLQDSLGGRTRTSTIATISPASLNLEE 345  
QY 346 TLTSTLEYAHRANKNLNKPVPNQKLTKKALKEYTTEBIEBKRLDAAAREKNG----- 396  
Db 346 TLTSTLEYAHRANKNLNKPVPNQKLTKKALKEYTTEBIEBKRLDAAAREKNG----- 396

RESULT 6

US-09-723-153-2  
; Sequence 2, Application US/09723153  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1048  
; CURRENT APPLICATION NUMBER: US/09/723,153  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 09/634,957  
; PRIOR FILING DATE: 2000-08-06





```
Db 2204 VAVLEEKLELENLQIISOQKGELBWKOLLERDKRTERMTAESRALQSCVECLSKKE 2263
QY 454 -----TWEEKLAKT-----EETHKREAAALGELGISIEK-----482
Db 2264 DLQEKDIEWKKLAOTKRVLAEEENSMEQSNLEKLELVKRVKLQQLDLNROKLSLHN 2323
QY 483 -----GFVGP-----YHSEKEMPHLVNLSDDPPLACIYINI 513
Db 2324 DISAMQOQLQOLVDLTGLATITSSKDGITSSFFHLFFHLKOLLHTTK-HQDVLLSE-----2376
QY 514 KPGQTRV-----GNVQDTPQ-----EIRLNGSKILKEHCTFFENVNV 552
Db 2377 ---OTRLQKDISWANREFCOKEETKQOQLVQLONEIENKLVQOEMNFORLQKE- 2432
QY 553 TIVPNKAAVVNCVRI-----DKPTRLRSGYRIILGDHFIFRHNHPEARAE 600
Db 2433 --RESESKLTSKVTLKEQHOQLKEITDQSKDLSKVLAAEERVITLQEEENWCE 2490
QY 601 RQBSLRLHSVTSNOLGSPAPGRHRTLSKAGSDADGSDSDSPLPFHFGKDSDFYARR 660
Db 2491 SLEXTL---SOTKROLSE---REQOLVKSGLLALQKEADSMRADFSLLRNQFLTERK 2543
QY 661 EASATILGIDQKI---SHLTDDLDLAFDD---VOKARAVRGLVEONED-----S 705
Db 2544 KAERQVASLKAIKIORSLEKNLLEORQNSCIQKEMATIELVAQDNHERARLMKELN 2603
QY 706 DSQSSFPVRKYSNCTIDNPSLDTAITMPTGPRSDDDGDALFFGDKKSKODASNVDFEE 765
Db 2604 QMOTXYTELKQMAN-----SOTKROLSE---REQOLVKSGLLALQKEADSMRADFSLLRNQFLTERK 2543
QY 766 LRQOQAQMEELAKTAKQE 783
Db 2622 LERRQMEISDAMRTLKASE 2639

RESULT 10
US-60-248-505-1112
; Sequence 1112, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000916
; CURRENT APPLICATION NUMBER: US/60/248,505
; NUMBER OF SEQ ID NOS: 1996
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 1112
; LENGTH: 992
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-1112

Query Match 3 28; Score 130; DB 6; Length 992;
Best Local Similarity 19.08; Pred. No. 0.076;
Matches 149; Conservative 128; Mismatches 286; Indels 222; Gaps 35;

QY 85 FQD-LGVPL---LDNAFKGYNKCIFAYQSGSKSYNMWYKKEHGV-----IP----R 130
Db 55 FQDSQGVPLPGTRNTCTFCGYDHLTSTFCRLS-SSAKWSGICGNKRAAGEPGTSHPPPKKA 113
QY 131 ICODMFRINLELO-KDKNLCTVEVSYLEIYNERVDLNPSTCKGNLKYREH---PSTGP 186
Db 114 AVEDSGTVEYIKLGVSGSTEELDITRLQYKNNKLAEMLDQRAQLEDELRHEIKLERQ 173
QY 187 VYEDLAKLVRSF--QETENL-----NDQC-----NKARTVAANTNHEISSHSA 229
Db 174 ATDDASILIVNWSQFENIRILKLYDLEQGLDGLLTERKALVVEPEPDSNGE-- 231
QY 230 VFTLTQKHDEETKMDTEKVAKISLVLDLAGSERATGTGATGARLKEAGAEINSLTIG 289

Db 232 -----RKDRNGKWEQBPAPFSLATLASS-----SSEMESQLOERVESSR-----274
QY 290 RVYIAADASSGQKGNOLVPRDSVLTWLLKDSLGGNSMTAMIAISPADINEFTLST 349
Db 275 RAVSQVTYVDKLQEKVELLS-----RKLSGCDNLIV-----BEAVQE 312
QY 350 LRYADSAKRINKHAVVNDPNRMATRELKE---LAQLRSKLQSGGGGGAGGSGGPV 405
Db 313 L-----NSFLAQENMRLOELTDLLOEKHRTMSQEFSLQSK-----V 349
QY 406 EESYPTDTPLEKQIVSIQDPDATVKKMSK-----AEIVPOLN-OSEKLYRDLNNTWEEK 458
Db 350 ETASRVSVLESMIDDLQWDIDKIRKREQRLNRLHAEVLKRVNSGYKYVGAGSLYGGT 409
QY 459 LAKTEIHKREAAALEEGISIEKPGVPGPHYSKEMPHLVNLSDDPPLACIYINIKPGQT 518
Db 410 ITINARKPFENNALEE-----NKE-----LAQNRL-----CELEKLRQDFE 446
QY 519 RVGVNQDTPQAEIRLNGSKILKEHCTFFENVNVITVPNE-----KAAVVNVNVR- 568
Db 447 EYTTONEKLVKELRSVAVQVVKETPEYRCMQSOFSVLYNESLQKLAHLDEARTLLRTRG 506
QY 569 -----IDKPTR-----LRSGYRIILGDHFIFRHNHPEARAEARQBSL-- 606
Db 507 THQHOVELIERDEYSLHKKLRTETVQLQEDTLAQRVKEVEMLRLEFEOTLAAEQAGPINR 566
QY 607 -LRHSVT-----NSQLGSPAPGRHRTLSKAGSDADGSDSDSPLPFHFGKDSDFYARR 660
Db 567 EMRHLSSLQHNHQLKGEVL-RYKRLRAQSDLN-----RTRL 605
QY 661 EAASAILGIDQKISHLTDD---ELDALFDYQKARAVRGLVED-NEDSDQSFPF---V 713
Db 606 RSGSALQSQSTEDPKPEALKPDSEDLSSQSSASKASQADANEIKKAQSKQEMKLL 665
QY 714 RDKYNSNGTIDNFSLOTATMPTGPRSDDDGDALFFGDKKSKODASNVDFEELRQOAOH 773
Db 666 LDMYRS-----APKEQRDKVOLMAAEKSK-----AELEDRLORLKD 703
QY 774 EALK 778
Db 704 EDKEK 708

RESULT 11
US-60-248-505-895
; Sequence 895, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000916
; CURRENT APPLICATION NUMBER: US/60/248,505
; NUMBER OF SEQ ID NOS: 1996
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 895
; LENGTH: 3078
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-895

Query Match 3 18; Score 126.5; DB 6; Length 3078;
Best Local Similarity 19.26; Pred. No. 0.66;
Matches 147; Conservative 115; Mismatches 252; Indels 253; Gaps 38;

QY 157 LEIYNRPVDLNPSTCKGNLKYREHSGVYEDLAKLVRSFOEINLMDGCKARTYA 216
Db 747 LAMWSKVKELCAKAEKLT---SHPSDAFOIEKEDLVSSWEHR-----790
QY 217 ATNNMETSRSRHAVFTLTQKH-----DEETKMDTEKVAKISL---VOLAGSE---263
```

Db 791 ATSRYE-----KLOATYWHRRSDFDELSCWNEKTAALINADELPTDVAGEVILL 841  
Qy 264 -----BATSTGATG-----ARLKGAEINSLSTGRVTAALADSSGQK 303  
Db 842 DRHQKHHEIDSYDRFQASDETGDVLNANHEASDEVREKWEILLONNWTALLLEWDERH 901  
Qy 304 KKNOL-----VYPRDS--VLTWL-----LKDSLGNSMTAMTAISPADINEFTSLT 350  
Db 902 ROYQOCLDFHLYRDSQVDSWMSRQEAFLNEEDLNSLSASAEALLQKE--DFEAFN-- 958  
Qy 351 RYASAK-----RIKNHAYVNEPNARKIRELKEALQSKLSQSSGGGGAGSGGPV 405  
Db 959 --AOEKIIVYDKTATKLIGDDHYUSENIKAIRUGLLARRDALREKA-----1003  
Qy 406 ESYVPTDPLEKQVISOQDPAVYKMSKAEIVQ-----LNQSEKL-----Y 448  
Db 1004 -----ATRRLLKESLQLVDESDLANINKKHLADDEDY 1042  
Qy 449 RDL-----NOTWEKLAETEEHKREAALEELGISTEKGVPGYHKSKEPHLVNI 499  
Db 1043 KDIQNLKSRVQKQVFEKELA---VNKTOLNTOKTGOEMIEG--GHYADNV--TTRL 1094  
Qy 500 SDPLLAELCLVYNIKPGQTRVGNVNDQTAETRLNGSKTLKEHCTFENVDNVVTVPNEX 559  
Db 1095 SEVASLWELLEATKQGTQLEHMQLOPE---NNAEDLQR--WLEDVEMQVTSYDCK 1149  
Qy 560 A-AMVNGVR---IDKPRLSRGYRIILGDPHF--RHNPE--EAAERQESLRLHS 610  
Db 1150 GLAEVNRKRIKGLLESVAARQDQVDLTDLAAYFEELGHPDQSKDIRA--RQESLVCRFE 1208  
Qy 611 VTNQSLGSPAGR-----HDKTLKSGADG--DSKSDSPULPHFRGX-----652  
Db 1209 A-----LKEPLATRKKLILLQILQICLDEDEMIQETESATSTLGLKLLASKLL 1264  
Qy 653 -----SDMYARBEAASAILGLDOKTISHLTDDELALPD 686  
Db 1265 NHRVINTASIEPIQIBTERGNKVPGEHFAEDVASYKSLKQNMESL-----1314  
Qy 687 DVOKARAVRGLVEDNEDSDQSFP-----VRDKVNSGNTINFSLDTATM 734  
Db 1315 --RABAARR-----ONDLEANVQOQVYLADLHEATWIREK---EPVDNTNGADEEA 1363  
Qy 735 PTPRSDDG---DALFFGDX--KSKDASNVVDELRQQAQMEAL 777  
Db 1364 AGALLKKEAFLLDUNSGFSGMALKRQNA-----CQOQQAAPVEGV 1406  
RESULT 12  
US-09-739-449-10081  
; Sequence 10081, Application US/09739449  
; GENERAL INFORMATION:  
; ORGANISM: Agrobacterium tumefaciens  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 10081  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-10081

Query Match 3.0% Score 122.5; DB 5; Length 619;  
Best Local Similarity 20.0% Pred. No. 0.13;  
Matches 136; Conservative 90; Mismatches 251; Indels 203; Gaps 32;  
Qy 78 YARO-EDLFDQGLDPLDNAPKYNICIFAYQGTGSKSY----SMWGYG-KHKGVTPT 131

Db 48 YQVQVSDLSKQFQTRITTEQKAKGIDQVFSOLAAGALDELAKSMWGLGSENRLANI 107  
Qy 132 COD-MFRFINELQKDKNLCTTVEVSYLETYNERVFDLLNPSTKNLKYREHSPGTPVED 190  
Db 108 AEDPAKSYNG-QFDNL-----FSEKLR-----NSGFRE-----DYIKE 142  
Qy 191 LAKLVRSFOEINILMDSGHKARTVAATMMNETSSRSHAVFTL-----T 234  
Db 143 RSKVAVRS--QIVAVSDGFAPOVLVDALKQYRQEARVDVILSNVAVPPVPAKPDV 200  
Qy 235 LTKQKHDEETKMDTEKVAKLSVLDLAGSERATSGATGARKLEGAEINR--SLSTLGR--- 290  
Db 260 LTPWFETNKSRYAPPEFKFTFYVKLPEPSDIAEPSVTDATDYNSHKSDSPRTAGRTV 260  
Qy 291 -----VTALADMSGQKKNQILVYPR-----DSVLTWLLKSLGNSMTAMIAA 335  
Db 261 EQITPEPKMAAAARAEQIRUGNTYDQVKKDGKTASVDTLGEFTADTIPQSI-----314  
Qy 336 ISPADINFEETLSTLAYASAKRIKKNHAYVNEPNARKIRELKEALQSKLSQSSGGG 395  
Db 315 -----ADAAFAI-----QKDGYS 328  
Qy 396 GGAGSGGPV---EESYPPDT--PLEKQVISOQDPAVYKMSKAEIVEQLNQSEKLYRD 450  
Db 329 PVYDGSFGPILLRVTCIKAPETTRTLDKAKEDIR-----KDLATMAAAEBITNVHDSYED 382  
Qy 451 LN-----OTWEKLAETEEHKREAALEELGISTEKGVPGYHKSKEPHLVNLSDDP 503  
Db 383 LRAGSSSLADAOKLMLKPVTDADAGLNEKGDALF--GLPSP-----OLAQD-- 430  
Qy 504 LLAELCLVYNIKPGQTRVGNVNDQTAETRLNGSKTL--KEHCTFENVDNVTV--IYVNEK 559  
Db 431 -----VFTPEG--TDALPTLGRGTYWFDVQIIPAKDRTLAERVDDVVDVADMTABOR 483  
Qy 560 AMVNGVR---VRDKPRLSRGYRIILGDPHFIRPNHPEEARERQESLRLHS-----610  
Db 484 TALAKADELAARVEKGETLEA---VAGELSL-----AVEDKSLGRRTSEDAIF 529  
Qy 611 --VTNSQLSGPAPRIHRTLSKSGADGDSR-----SDSPLPHFRGKSDSNFYAR 659  
Db 530 GRTTAAVESAEG-----VVGTAADADGSSRLFKVTSVDNANAPALANDQOQFATA 584  
Qy 660 REAASATLGLDQKISHLTD 679  
Db 585 RAAGDDM--LDQWVNLQND 602  
RESULT 13  
US-09-739-449-9117  
; Sequence 9117, Application US/09739449  
; GENERAL INFORMATION:  
; ORGANISM: Agrobacterium tumefaciens  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15490)C  
; CURRENT APPLICATION NUMBER: US/09/739,449  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 09/514,000  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 13351  
; SEQ ID NO 9117  
; LENGTH: 1108  
; TYPE: PRT  
; ORGANISM: Agrobacterium tumefaciens  
US-09-739-449-9117

Query Match 3.0% Score 121.5; DB 5; Length 1108;  
Best Local Similarity 18.9% Pred. No. 0.34;  
Matches 145; Conservative 121; Mismatches 273; Indels 229; Gaps 37;  
Qy 114 KSYSMWGYCKHKGVT-----IPRLQCFMRINELQKDKNLCTTVEVSYLETYNE-----RV 164

```

Db 214 KAQGMENAKOOGTASIKLIPELREDEARVAALOR-----LQIARTQLDDEANRLRR 266
QY 165 RDLL--NPSTGNLKVREHSTGPVYEDLAKLVRSFOETENIM-----DEGKA 212
Db 267 RDELARLSOLGDIVREE---RLVADNAQILARLDEEAELLDLSDSGRHADENREA 322
QY 213 RTVAATNMETSSSHAVFTLTUTOK-----WHDEETKMD---TEKVAK 253
Db 323 FEMAAVKLAE---SEAVFTSITAEAAAGQQLERAIRDLSDRKLRLEQRQSEASAE 378
QY 254 ISLVD--LAG---SERATSTGATGARKEG---AFINRSLSLTGLRVIAALA--DMSSCK 302
Db 379 IDTIDELKSLGDPFAERREAVEAEIATEDALIVAEAEAAVAEARSAAELAGPLETAK 438
QY 303 OKKNOLVPPYRDSVLTWLLKDSLGSNMTAAIAPADINFEETLSLTRVADSARKIKNI 362
Db 439 NFLNAL--DTEARTITKMLATSAANGSFTPVAEMTVBERGYEALG-----483
QY 363 AVYNEDPNARMIRELKEELAOLRSKIQSS-----GGGGGAGSGGPPVESYPPDTPLEK 417
Db 484 AALGDD-----LESPDASAPAYWGGNGGADDPGLPOGAK-----PL--521
QY 418 QIVSIQOOPATVKMKAIEIVEOINOSKLYRLNO-----TWEELAKTE--463
Db 522 --LDYQAQDALLRALAQIGTVADVSEARLLPSIKAGORLVTREGALFRMDGHIASADAP 580
QY 464 ---EIKHEAALEELGISTGIEKGVGYHSEMPHLVNLSDPPLAELCLVYIKRGQTRV 520
Db 581 GAAALRSOKNRLAEIAEELD-----EARSILEAEADQLAAK--TEDIRSELR 628
QY 521 GNVNQD-----TQAEIRLNGSKILKEHCTFENVNVNVTIYPNEKAAVMYN 565
Db 629 SEVDRASRLATROLAAREALITSAE--RASGDLRLRRDVSEAOINQAGIT--DEIAYQEE 685
QY 566 GYRI---DKP-----TRLRSYRIILGDFHIFRNHPEAREARQOSLLRHVSNSOL 616
Db 686 NARTEMEDAPLSYDLRLRRESOLEVATDGLL-----AARAREGVIS--REAESKOR 737
QY 617 GSPAPGRHDT--LSKAGSDADGSDRSPLPHFRGKSDMFYARREASAILGLDQKISH 675
Db 738 RIQAIGOEKSTWASRAASAD-----HIATLUREEE---ARETAEIDIAPEE---783
QY 676 LFDDELDAFDQKARAVRGLVEDNEDSDOSSFPVRKQYKMSNGTIDNFSLDATIMP 735
Db 784 --FDEKRRNLLNEIQKTEDARAAADRLAEANLQIR--AAORVAA-----TALS--828
QY 736 GTPRSDDGDGALFPGDKSKODASVNDVEELRQOQAQMEALAKTAKOE 783
Db 829 -----ELAEAREKRGAEERLVSAREK 850

RESULT 14
US-08-945-567C-4
; Sequence 4, Application US/08945567C
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSEMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KIM, Michael H.
; TITLE OF INVENTION: MOXAXELLA
; FILE REFERENCE: 1038-7/05 MTS
; CURRENT APPLICATION NUMBER: US/08/945, 567C
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431, 718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478, 370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621, 944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
```

```

; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1833
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
; US-08-945-567C-4

Query Watch          3 08; Score 120; DB 4; Length 1833;
Best Local Similarity 19 48; Pred No 0 88;
Matches 160; Conservative 116; Mismatches 295; Indels 254; Gaps 46;

QY 5 GNIVVVVRVRFNAREIDRGAKCIYRMSGNOTILTP--PPGAEEKAR--KSKGKTIMOGPK 60
Db 560 GNVKI-----TNTGINAGKAI-----TGLSPTLPSIADQSSRIELGNTIQ-----601
QY 61 AFAPDRSYWSPDKNAPNARCEDLFDGLVPLLDNAF--KGYNNCI--PAYGOTGSGKSY 117
Db 602 -----DKDSNAASIND-----ILNTGFNLKNNNPIDFV-----STYD 635
QY 118 MMGYGKEHGVIPRICODFPRINELQOKNL--TCTVEVSYLEIYNERVOLLNPSTGNL 176
Db 636 IVDFAANGNATTATVTHDTANTKSKVYDVNVDDTIHL-----LNTGFPNLKNNNPIDFV-----673
QY 177 KYREHPSTGPVYEDLAKLVRSFOETENIMDEGNKARTVAATNMNETSSSHAVFTL--233
Db 674 -----TG--TDDNKKLGCVT--TKLNTKSANGN-----TATNFVNSSDEDALVNAKDI 718
QY 234 -----TITQKWH-----DEETKMDTEKVAKISLVDLAGSERATSTGATGAR-----LKE 277
Db 719 ABENLNTLAKELHTTKTADTALQTFYKKVDENNADANAIVTGOKNANQVNTLTGK 778
QY 278 GAENIKSLSLGRVIAALADMSGOKKKNQOLVPYRDSVLTWLLKDSLGSNMTAMIAIS 337
Db 779 ENGLNKTDKNGTGVGTGINTTSLGKAGKSTL---NDGGLS--TKNP7GSSQIQV-----827
QY 338 PADINFEETLSLTRVADSAR--RIKNHVAVNE--DPNARMIRELKEELAOLRSKLOSGG 393
Db 828 -----GADGVKFAVNNNGVAGIDGTTTRITR---DEI-----858
QY 394 GGGGAGSGGPPVESYPPDTPLEKQIVSIQOOPATVKMKAIEIVEOLNOS---EKLYRD 450
Db 859 ---GFTGTNGSLDKSP---HLSKD--GINAGGKKITNIQSGEIAQNSHDAVTGGKIY-D 909
QY 451 LNQTWEEKLAKTEEIIHKEAREAALEELGISIEKG---FVGPHYHSEMPHLVNLSDPPLA 506
Db 910 LKTELEKNTSSSTA---KTAQNSLHSEFVADQGNNTFVSNPYSSYDTS---KTSDVITFA 963
QY 507 ---ECLVYNIKFGQTRVGNVQD---TQAEIRL---NGSKILKE---HCTFENVNVNVTI 554
Db 964 GENGITTKVNGGVVRVG-IDQTGLTPTKLYGVNNNGKIVIDSONGONTITGLSNTLAN 1022
QY 555 VNEKAAVMY-----NCVRIDKPTRLSRGYRIILGDFHI-----FRFNH 593
Db 1023 VTNDKGSVTEQGNIIKDEKTRAASIVDVLGAFNGLOGNSEAVDFVSTYDYNFADGN 1082
QY 594 PEAREAROBOSLLRHISV-----TNSQLGSPAPGRHDTLSKAGSDADGSDRSPLP 646
Db 1083 ATTAKVYTDTSKTSKTVYDVNVDDTTEVADKKLGVKTTTLTSTGTGANKFALSN----1138
QY 647 HFRGKSDMFYARREASAILGLDQKISHLTDDDELDAFDQKARAVRGLVEDNEDS-705
Db 1139 -----QATGDALVKASDIYAH-----LNTLSGDIQTA---KGASQANNSAG 1176
QY 706 ---DSQGSFFVRD---KY---MSGNTID---NFSLDATIMPCTP 738
Db 1177 VYDAGNKVYIDSTDNKNIYQAKNDGIVDKTEKVAKDKLVAAQAQTP 1221

RESULT 15
US-08-945-567C-3
; Sequence 3, Application US/08945567C
```



```

: GENERAL INFORMATION:
: APPLICANT: SASAKI, Ken
: APPLICANT: HARKNESS, Robin E.
: APPLICANT: LOOMORE, Sheena M.
: APPLICANT: CHONG, Pele
: APPLICANT: KLEIN, Michael H.
: APPLICANT: LEWIS, HIGBAXELLA
: TITLE OF INVENTION: HIGBAXELLA
: FILE REFERENCE: 1038-745.MIS
: CURRENT APPLICATION NUMBER: US/08/945,567C
: CURRENT FILING DATE: 1998-03-19
: PRIOR APPLICATION NUMBER: 08/431,718
: PRIOR FILING DATE: 1995-05-01
: PRIOR APPLICATION NUMBER: 08/478,370
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 08/621,944
: PRIOR FILING DATE: 1996-03-26
: PRIOR APPLICATION NUMBER: PCT/CA96/00264
: PRIOR FILING DATE: 1996-04-29
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO: 3
: LENGTH: 1992
: TYPE: PRT
: ORGANISM: Moraxella catarrhalis
: US-08-945-567C-3

```

```

Query Match          3.0%; Score 120; DB 4; Length 1992;
Best Local Similarity 19.4%; Pred. No. 0.99;
Matches 160; Conservative 116; Mismatches 295; Indels 254; Gaps 46;

QY 5 GRKVVYVRFPFNAREIDRGAKIYRMEGNQITLP--PGAEAKAR--KSGKTMIDGPK 60
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 719 GNWKI-----TWNGINAGKAI-----TGLSPTLPSIADOSSRIELGNTIQ----- 760

QY 61 AFADRSWGFQKAPNARQEDLDQGLVDNAP--KGYNCCI--PAYCOTGSGKSYIS 117
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 761 -----DROKSNAASIND-----LNTGFLKNNNPIDFV-----STYD 794

QY 118 MWGYCKEHGVIPIRCODMFRINELQDKKL--TCTVEVSYLEIYNEVRDLAPSTKGNL 176
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 795 IVDPANGNATTATVHTDANKTSKVYDVNVDDTTIHL----- 832

QY 177 KYREHFTSGPYVEDLAKLVKRSFQBIENLMDGSKARTVAATNMNYSRSRSHAVFTL--- 233
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 833 -----TG--TDONKKLGVKT--TKLTKTSANGH-----TATNFNVNSSDEDALVNAKDI 877

QY 234 -----TLTQKWH-----DEETKMDTEKYAKISLVDLAGSERATSTGATGAR-----LKE 277
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 878 AENLNTLAKELIHTTKTADTALQFTYVKYVDENNADANAIVGQRKANNQVNTLTLAG 937

QY 278 GAEINSLSTLGRVIAALADNSSGQKKQNLVYPRDSVLTWLLKDSLGGNSMTAMIAIS 337
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 938 ENGLINKTDNGTVFTGINTISGLAKGKSL---NDGGLS--LRNPTSGEIQV----- 986

QY 338 PADINEETLSTLRVADSAAK--RIKNHAYVNE--DPNARMIRBELKELAQLRSLQSSGG 393
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 987 -----GADGVKFAKVNNGVYVAGIDGTTRITR---DEI----- 1017

QY 394 GGGAGSGGPGVEESYPPDPLEKQIVSTQOPDATYKMKSAEIVLQLNOS---EKLYRD 450
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1018 ---GFTGTNGSLDKSKP---HLSKO--GINAGCKKTTMIQSGEIAQNSHDVATGGKIY-D 1068

QY 451 LNOTMEKELAKTEBIHKEREAALEELGISIEKG---FVGPHYSKEMPHLYNLSDDPILA 506
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1069 LATELENKISSTA---KTAQNSLJHEFSAVADQGNFTVSNPYSSYDTS---KTSQVITEA 1122

QY 507 --ECLVYHAKPQOTRYGVNQD---TQAEIRL---NSGKILKE-----HCTFEWQNVVTI 554
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1123 GENGITKVKGVKGVYRG--IDQTVGLTTPRLPYGVNNGKGVIDSONGQNTITGLSNTLAN 1181

QY 555 VPNEKAAMVW----NGVRIDKPTRLRSRGYRIILGDPHI-----FRFNE 593

```

```

Db 1182 VTNDKGSVTTTEQGNIIKDEKTRAASIVDVLVSAGFNLOGNGEAVDFVSTYDTVNFADGN 1241
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 594 PEAAARQEQSLLRHSV-----TNSQLGSPAFGRHDIRTLKSGASDADGDSRSDSPUP 646
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1242 ATTAKVTYDDTSKTSKVYDVNVDDTTIEVKDKKLGKVTTLTSTGTGANKFALSN---- 1297
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 647 HFEGKDSWTFARRRASAAILGLDQKISHULTDDELDALPDYQKARVRGRLGVEDNEDS- 705
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1298 -----QATGDALVYKASDIVAH-----LNTLSGDIQTA-----KGASQANNSAG 1335
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 706 --DSQSSFPVRD---KV---MSNGTID---NFSLDTAITMPGTP 738
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 1336 YVDAGNKNVIYDSTDNKYQAKNDGTVDKTEKAVANDKLVAQAQTP 1380
   ||:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

Search completed: April 25, 2001, 10:16:29  
Job time: 300 sec





LOCATION: (443)..(601)  
OTHER INFORMATION: stalk domain, unc-104 family domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (602)..(784)  
OTHER INFORMATION: tail domain  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (713)  
OTHER INFORMATION: polymorphic variant #1 Val -> Ile  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (762)  
OTHER INFORMATION: polymorphic variant #2 Asp -> Glu  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (774)  
OTHER INFORMATION: polymorphic variant #3 Glu -> Asp  
PCT-US99-01355-1

Query Match 100.0%; Score 1834; DB 1; Length 784;  
Best Local Similarity 100.0%; Pred. No. 1.3e-173;  
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGGNIKVVYVRPNAREIDRGAKCIVRMEGNOTILTPPGAEKARKSGKTIMDGP 60  
DB 1 MSGGGNIKVVYVRPNAREIDRGAKCIVRMEGNOTILTPPGAEKARKSGKTIMDGP 60  
QY 61 AFAPDASYWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFYCGTSGSKSYSMG 120  
DB 61 AFAPDASYWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFYCGTSGSKSYSMG 120  
QY 121 YGKEHGVIPTICODMFRRINLQKDKNLTCTVEVSYLEYINERVRDLNPNSTGKMLK 180  
DB 121 YGKEHGVIPTICODMFRRINLQKDKNLTCTVEVSYLEYINERVRDLNPNSTGKMLK 180  
QY 181 HPSTGYPVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKH 240  
DB 181 HPSTGYPVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKH 240  
QY 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTGKRVIAALADMS 300  
DB 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTGKRVIAALADMS 300  
QY 301 GQKKKQLVPYRDSVLTWLLKDSLGNSWMTAAISPADINFEETLSTLRYADSAK 357  
DB 301 GQKKKQLVPYRDSVLTWLLKDSLGNSWMTAAISPADINFEETLSTLRYADSAK 357

RESULT 2  
US-09-235-416-1  
Sequence 1; Application US/09235416A  
GENERAL INFORMATION:  
APPLICANT: Sakowitz, Roman  
APPLICANT: Goldstein, Lawrence S. B.  
TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor  
TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor  
FILE REFERENCE: 1857C-000710US  
CURRENT APPLICATION NUMBER: US/09/235,416A  
CURRENT FILING DATE: 1999-01-22  
EARLIER APPLICATION NUMBER: WO PCT/US99/01355  
EARLIER FILING DATE: 1999-01-22  
EARLIER APPLICATION NUMBER: US 60/072,361  
EARLIER FILING DATE: 1998-01-23  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 784  
TYPE: PRT  
ORGANISM: Thermomyces lanuginosus  
FEATURE:

OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed  
OTHER INFORMATION: microtubule motor protein  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (1)..(357)  
OTHER INFORMATION: kinesin-like microtubule motor domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (358)..(442)  
OTHER INFORMATION: neck domain links motor domain to stalk domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (443)..(601)  
OTHER INFORMATION: stalk domain, unc-104 family domain  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (602)..(784)  
OTHER INFORMATION: tail domain  
US-09-235-416-1

Query Match 100.0%; Score 1834; DB 16; Length 784;  
Best Local Similarity 100.0%; Pred. No. 1.3e-173;  
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGGNIKVVYVRPNAREIDRGAKCIVRMEGNOTILTPPGAEKARKSGKTIMDGP 60  
DB 1 MSGGGNIKVVYVRPNAREIDRGAKCIVRMEGNOTILTPPGAEKARKSGKTIMDGP 60  
QY 61 AFAPDASYWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFYCGTSGSKSYSMG 120  
DB 61 AFAPDASYWSFDKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFYCGTSGSKSYSMG 120  
QY 121 YGKEHGVIPTICODMFRRINLQKDKNLTCTVEVSYLEYINERVRDLNPNSTGKMLK 180  
DB 121 YGKEHGVIPTICODMFRRINLQKDKNLTCTVEVSYLEYINERVRDLNPNSTGKMLK 180  
QY 181 HPSTGYPVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKH 240  
DB 181 HPSTGYPVEDLAKLVRSFQEIENLMDGKNKARTVAATNMNETSSRSRHAVFTLTQKH 240  
QY 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTGKRVIAALADMS 300  
DB 241 DEETKMDTEKVAKISLVDLAGSERATSTGATGARLKEGAEINRSLSLTGKRVIAALADMS 300  
QY 301 GQKKKQLVPYRDSVLTWLLKDSLGNSWMTAAISPADINFEETLSTLRYADSAK 357  
DB 301 GQKKKQLVPYRDSVLTWLLKDSLGNSWMTAAISPADINFEETLSTLRYADSAK 357

RESULT 3  
US-09-654-850-1  
Sequence 1; Application US/09654850  
GENERAL INFORMATION:  
APPLICANT: Sakowitz, Roman  
APPLICANT: Goldstein, Lawrence S. B.  
TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor  
TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor  
FILE REFERENCE: 1857C-000710US  
CURRENT APPLICATION NUMBER: US/09/654,850  
CURRENT FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 09/235,416  
PRIOR FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: US 60/072,361  
PRIOR FILING DATE: 1998-01-23  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 784  
TYPE: PRT  
ORGANISM: Thermomyces lanuginosus  
FEATURE:

```

; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; NAME/KEY: DOMAIN
; LOCATION: (358)..(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; NAME/KEY: DOMAIN
; LOCATION: (443)..(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; NAME/KEY: DOMAIN
; LOCATION: (602)..(784)
; OTHER INFORMATION: tail domain
US-09-454-850-1

Query Match      100.0%; Score 1834; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.3e-173;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGGNIKVVYRPFNAEIDRGAKCIVRMGGNOTILTPPGAEEKARKSGKTINDGPK 60
DB 1 MSGGGNIKVVYRPFNAEIDRGAKCIVRMGGNOTILTPPGAEEKARKSGKTINDGPK 60
QY 61 AFADRSWGSFQKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGTGSGSKYSMMG 120
DB 61 AFADRSWGSFQKNAPYARQEDLFQDLGVPLLDNAFGYNNCIFAYGTGSGSKYSMMG 120
QY 121 YGKEHGVPICQDWFREINELQKDKNLTCTVEVSYLEIYNERVDDLNPSTKGNLKYRE 180
DB 121 YGKEHGVPICQDWFREINELQKDKNLTCTVEVSYLEIYNERVDDLNPSTKGNLKYRE 180
QY 181 HPSTGPGVEDAKLWRSQEIENLMDGKNARTVAATNNNETSSRSRSHAVFTLITQKWH 240
DB 181 HPSTGPGVEDAKLWRSQEIENLMDGKNARTVAATNNNETSSRSRSHAVFTLITQKWH 240
QY 241 DEETKMDTEKVAKISLVLDLAGSERATSGATGARLKEGAEINRSISLTGLRVIAALADSS 300
DB 241 DEETKMDTEKVAKISLVLDLAGSERATSGATGARLKEGAEINRSISLTGLRVIAALADSS 300
QY 301 GQKKKQLVPPRDSVZVLLKSLGSGNSMTAMTAISPADINFEETLSTLRYADSAK 357
DB 301 GQKKKQLVPPRDSVZVLLKSLGSGNSMTAMTAISPADINFEETLSTLRYADSAK 357

RESULT 4
US-09-417-507-37298
; Sequence 37298, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 37298
; LENGTH: 533
; TYPE: PRT
; ORGANISM: A. fumigatus
; NAME/KEY: UNSURE
; LOCATION: (515)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-417-507-37298

Query Match      83.3%; Score 1528.5; DB 18; Length 533;
Best Local Similarity 86.7%; Pred. No. 2.1e-143;
Matches 294; Conservative 27; Mismatches 17; Indels 1; Gaps 1;

QY 20 EIDRGAKCIVRMGGNOTILTPPGAEEKARKS-GKTINDGPKAFAFADRSYNSFQKNAPY 78

```

```

DB 8 EIDRGAKCIVOMKNGQTLVPPGADSKSRKAGGAGVEGPKTFAFDSYSWFSQKNAPY 67
QY 79 ARQEDLFQDLGVPLLDNAFGYNNCIFAYGTGSGSKYSMMGYGKERGVIPRICQDWFRR 138
DB 68 AGODMLFADLVPLLDNAFGYNNCIFAYGTGSGSKYSMMGYGKERGVIPRICQDWFRR 127
QY 139 INSLQKOKNLCTCTVEVSYLEIYNERVDDLNPSTKGNLKYREHREFTGPGVEDAKLWRS 198
DB 128 IAKMDEKLNCTCTVEVSYLEIYNERVDDLNPSTKGNLKYREHREFTGPGVEDAKLWRS 187
QY 199 PQEIENLMDGKNARTVAATNNNETSSRSRSHAVFTLITQKWHDEETKMDTEKVAKISLV 256
DB 188 FEIDRLMDGKNARTVAATNNNETSSRSRSHAVFTLITQKWHDEETKMDTEKVAKISLV 247
QY 259 LAGSERATSGATGARLKEGAEINRSISLTGLRVIAALADSSGYGKKKQLVPPRDSVZV 318
DB 248 LAGSERANSTGATGARLKEGAEINRSISLTGLRVIAALADVAGSKKKKNSAVPYRDSILTW 307
QY 319 LKDSLGLSGNSMTAMTAISPADINFEETLSTLRYADSAK 357
DB 308 LKDSLGLSGNSMTAMTAISPADINFEETLSTLRYADSAK 346

RESULT 5
US-09-488-725A-6751
; Sequence 6751, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: PCT/US98/000001
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-15
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL-Genes_b Versions 1.0
; SEQ ID NO 6751
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-6751

Query Match      62.8%; Score 1152; DB 18; Length 421;
Best Local Similarity 62.1%; Pred. No. 5.4e-106;
Matches 226; Conservative 58; Mismatches 56; Indels 24; Gaps 6;

QY 4 GGNIKVVYRPFNAEIDRGAKCIVRMGGNOTILTPPGAEEKARKSGKTINDGPKAFA 63
DB 9 GASVYKAVYRPFNRSKSGKCIQNSGSGSTTIVNPKQPKET-----PRFSF 57
QY 64 FDSYSWFSQKNAP---NYARQEDLFQDLGVPLLDNAFGYNNCIFAYGTGSGSKYSMMG 120
DB 56 FDSYSWNS--HTSPEDINASQKVYRDLGEENLQHRFEGINCFAYGTGSGSKYSMMG 115
QY 121 YGK--BHGVPICQDWFREINELQKDKNLTCTVEVSYLEIYNERVDDLNPSTKGNLKY 178
DB 116 KQEKDQGGILPOLCEDLFSRINDTND-NMSYSVEYSYMEIYCERVDDLNPSTKGNLKY 174

```

```
QY 179 REHPSTGPVEDLAKLVRSFOEINLMDGKNCARTVAATNMNETSSRSYAVFTLTITOK 238
DB 175 REHPPLGPPVEDLSKLVTSYNDIQDMDGKNCARTVAATNMNETSSRSYAVFTITOK 234
QY 239 WHDEETKMDTEKVAKISLVLDLAGSERATSGATGARKLKEGAETNLSLTGLRVTAALADM 298
DB 235 RHDAETNITTEKYSKISLVLDLAGSERADSTGATGARKLKEGANINLSLTGLKVISALAE 294
QY 299 SSG-----KQKNOLVPRDSVLTWLLKDSIGGNSWTAMTAALSPADINFEETSLTRYA 353
DB 295 DSGPNKMKKKKTDFTPYRDSVLTWLLRNLGNSRTAMVAALSPADINDETSLTRYA 354
QY 354 DSAK 357
DB 355 DRAK 358

RESULT 6
US-60-258-275-444
; Sequence 444, Application US/60258275
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001026-PROV
; CURRENT APPLICATION NUMBER: US/60/258,275
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 717
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 444
; LENGTH: 1793
; TYPE: PRT
; ORGANISM: HUMAN
US-60-258-275-444
```

```
Query Match 62.8%; Score 1152; DB 23; Length 1793;
Best Local Similarity 62.1%; Pred. No. 5.6e-105;
Matches 226; Conservative 58; Mismatches 56; Indels 24; Gaps 6;

QY 4 GGNKVVVRPFNAREIDRGAKCIVRMEGNQTLITPPPGAEEKARKSGRTIMDGPKAFA 63
DB 9 GASVKAIVRPFNRSRMSKCIQMSGSTTIVNPKPKET-----PKSFS 57
QY 64 FDRSYWSFDKNAP---NYAREDLFODLGVPLLDNAFGYNNCIFYAGQTGSGKSYSMWG 120
DB 58 FDISYMS--HTSPEDINTASQKQYRDIGEMLQHAPEGYNVCIFYAGQTGAGKSYTMWG 115
QY 121 YGK--EHGVIPRICQDMFRINELQDKNLTCTVEVSLIYNRVDLLNPTGKGLKV 178
DB 116 KOEKQOGIIFQCELDLSRINTTND--NMSYVEVSYMEICERVADLLNPNKGNLKV 174
QY 179 REHPSTGPVEDLAKLVRSFOEINLMDGKNCARTVAATNMNETSSRSYAVFTLTITOK 238
DB 175 REHPPLGPPVEDLSKLVTSYNDIQDMDGKNCARTVAATNMNETSSRSYAVFTITOK 234
QY 239 WHDEETKMDTEKVAKISLVLDLAGSERATSGATGARKLKEGAETNLSLTGLRVTAALADM 298
DB 235 RHDAETNITTEKYSKISLVLDLAGSERADSTGATGARKLKEGANINLSLTGLKVISALAE 294
QY 299 SSG-----KQKNOLVPRDSVLTWLLKDSIGGNSWTAMTAALSPADINFEETSLTRYA 353
DB 295 DSGPNKMKKKKTDFTPYRDSVLTWLLRNLGNSRTAMVAALSPADINDETSLTRYA 354
QY 354 DSAK 357
DB 355 DRAK 358
```

RESULT 7

US-09-488-725A-3179

```
; Sequence 3179, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Huseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLOCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 3179
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-3179
```

```
Query Match 62.6%; Score 1149; DB 18; Length 893;
Best Local Similarity 61.8%; Pred. No. 3.6e-105;
Matches 225; Conservative 59; Mismatches 56; Indels 24; Gaps 6;

QY 4 GGNKVVVRPFNAREIDRGAKCIVRMEGNQTLITPPPGAEEKARKSGRTIMDGPKAFA 63
DB 3 GASVKAIVRPFNRSRMSKCIQMSGSTTIVNPKPKET-----PKSFS 51
QY 64 FDRSYWSFDKNAP---NYAREDLFODLGVPLLDNAFGYNNCIFYAGQTGSGKSYSMWG 120
DB 52 FDISYMS--HTSPEDINTASQKQYRDIGEMLQHAPEGYNVCIFYAGQTGAGKSYTMWG 109
QY 121 YGK--EHGVIPRICQDMFRINELQDKNLTCTVEVSLIYNRVDLLNPTGKGLKV 178
DB 110 KOEKQOGIIFQCELDLSRINTTND--NMSYVEVSYMEICERVADLLNPNKGNLKV 168
QY 179 REHPSTGPVEDLAKLVRSFOEINLMDGKNCARTVAATNMNETSSRSYAVFTLTITOK 238
DB 169 REHPPLGPPVEDLSKLVTSYNDIQDMDGKNCARTVAATNMNETSSRSYAVFTITOK 228
QY 239 WHDEETKMDTEKVAKISLVLDLAGSERATSGATGARKLKEGAETNLSLTGLRVTAALADM 298
DB 229 RHDAETNITTEKYSKISLVLDLAGSERADSTGATGARKLKEGANINLSLTGLKVISALAE 288
QY 299 SSG-----KQKNOLVPRDSVLTWLLKDSIGGNSWTAMTAALSPADINFEETSLTRYA 353
DB 289 DSGPNKMKKKKTDFTPYRDSVLTWLLRNLGNSRTAMVAALSPADINDETSLTRYA 348
QY 354 DSAK 357
DB 349 DRAK 352
```

RESULT 8

US-09-718-563-4

```
; Sequence 4, Application US/09718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
```

```

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-4

Query Match
Best Local Similarity 61.2%; Score 1123; DB 21; Length 348;
Pred. No. 3,1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKKVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
Db 2 GASVAVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
; SEQUENCE 4, Application US/09/718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-4

Query Match
Best Local Similarity 61.6%; Score 1123; DB 21; Length 348;
Pred. No. 3,1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKKVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
Db 2 GASVAVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
; SEQUENCE 4, Application US/09/718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-4

```

```

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-2

Query Match
Best Local Similarity 61.6%; Score 1123; DB 21; Length 348;
Pred. No. 3,1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKKVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
Db 2 GASVAVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
; SEQUENCE 2, Application US/09/718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-2

```

```

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-2

Query Match
Best Local Similarity 61.2%; Score 1123; DB 21; Length 348;
Pred. No. 3,1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKKVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
Db 2 GASVAVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
; SEQUENCE 4, Application US/09/718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-4

Query Match
Best Local Similarity 61.2%; Score 1123; DB 21; Length 348;
Pred. No. 3,1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKKVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
Db 2 GASVAVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
; SEQUENCE 4, Application US/09/718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,842
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-842-4

```

```

; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-2

Query Match
Best Local Similarity 61.6%; Score 1123; DB 21; Length 348;
Pred. No. 3,1e-103;
Matches 220; Conservative 59; Mismatches 62; Indels 16; Gaps 5;

Qy 4 GGNKKVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
Db 2 GASVAVVRVFPNAREIDRGAKIVRMGNGOTILTPPGAGAEKARKSGKTMGPKAFA 63
; SEQUENCE 2, Application US/09/718563
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718,563
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Human
US-09-718-563-2

```

Qy	4	GNIKYVVYRFPNARIDRGAKCIYVMBGQNTLLPDPGAEKAKRKSGKTYMDGPKAFA	63
Db	3	ASVAVYVYRFPNARIDRGAKCIYVMBGQNTLLPDPGAEKAKRKSGKTYMDGPKAFA	51
Qy	64	FDRSYWSF-DKNAENYARQEDLPDQGLVPLDINAKFYNNICFAYGQTGSGKSYSMGYG	122
Db	52	FQYSYWSHTSPDPCFASQNRVYNDIGKEMLLHAFEGYVNCIFAYGQTGAGKSYVTMMGQ	111

US-09-718-563-6  
; Sequence 6, Application US/09718563  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Freedman, Richard



```

: TITLE OF INVENTION: Novel motor proteins and methods for
:
: TITLE OF INVENTION: their use
:
: FILE REFERENCE: 1021
:
: CURRENT APPLICATION NUMBER: US/09/718,563
:
: CURRENT FILING DATE: 2000-11-22
:
: NUMBER OF SEQ ID NOS: 10
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 6
:
: LENGTH: 1770
:
: TYPE: PRT
:
: ORGANISM: Human
:
: US-09-718-563-6

```

Query Match	61.2%	Score 1123	DB 21	Length 1770	
Best Local Similarity	61.5%	Pred. No. 4-102			
Matches 220	Conservative	59	Mismatches 62	Indels 16	Gaps
QY	4	GGNIGKVVVRFPNAREIDGAKICTVMEGNOTILTPPTGCAEAKRKSGKTMQCPKPA	63		
DB	3	GASGVAAVRFPNRSRRESKSIQOMOGNSIINPKPK	51		-----APKFS
QY	64	FDRSWVSF-DKNAPYARQEDLPDQVLPDINAFKYNVCITFAYGTGGSKSYSMGVG	122		
DB	52	FDSYSWHSGSPEDCFASQNRVYNDIGKMLHSEGYNVCITFAYGTGGSKSYTMNGKQ	111		
QY	123	KEH--GVPIRCODMPRIEQLQDKRLCTVSEVSLYETINERVLDLNPSTGNLKVRE	180		
DB	112	ESQAGIITPOLCELEFXKIND--NCNEMSYSVSEVSEIYECRVDLLNPKNGNLVRE	170		
QY	181	HPSTGPVYEDLAKLVRSFQREIKMDQCNKARTVAATNMHSTSSSHAVFTLTQKWH	240		
DB	171	HPGLVYEDLSKLVFTLTADDAGNAGKARTVAATNMHSTSSSHAVFTVTFQKH	230		
QY	241	DEETKMTQTEKVAKTSIVDLGASGACGABKFGKGAETNSLIGRTVIAALADMS	300		
DB	231	DNETNLSTEKYSKTSIVDLGASRADSTGAGTRBKEGANINKSLTTLGKVISALAEYSK	290		
QY	301	GKQKKNOLVPRDSVITWLLKQSLGNSMTAATAISPADNEETLSLTVADSAK	357		
DB	291	-KKKKTDFIPVRDSVITWLLRENLGNSMTAAVAALSPADINDETLSLTVADRAK	346		

Search completed: April 25, 2001, 10:16:26  
Job time: 327 sec

```

RESULT      15
US-09-718-804-6
; Sequence 6, Application US/09718804
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; INVENTOR: Beraud, Christophe
; ATTORNEY: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1021
; CURRENT APPLICATION NUMBER: US/09/718, 804
; CURRENT FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1770
; TYPE: PRT
; ORGANISM: Human
US-09-718-804-6

```

	Query Match	61.2%	Score 1123	DB 21	Length 1770
Seqs Loca	Similarity	81.0%	Pred. No. 4.3e102		
Matches	Conservative	5%	Mismatches	62	Indels 16; Gaps
	220				
Qy	4	GNIVAVVVRPNARETDGACATVYBEGNQTLLTPPGAEAEKARSKCTITDGGPKAPA	63		
Db	3	GASVAVVVRPNARETSKESKCLIIOWGNSTSTNPKNKE	51		
Qy	64	FDRSYWVF-DKNAIPNARQEDFDQDGLVPLLDNAFKYKNCIFAYGQTSKGSVMKYG	122		

---

GenCore version 4.5  
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:16:29 ; Search time 7.94 seconds  
(without alignments)  
141.828 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_1\_357

Perfect score: 1834  
Sequence: 1 MSGGKGVVVRFPNARE.....PADINFEETLSTLYADSAK 357

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 17899 seqs, 3154390 residues

Total number of hits satisfying chosen parameters: 17899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_5/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_5/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_5/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_5/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_5/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_5/ptodata/1/paa/US06\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	612	33.4	368	5	US-09-724-519-10
2	612	33.4	382	4	Sequence 10, Appl
3	612	33.4	375	5	US-09-724-519-4
4	612	33.4	375	5	US-09-724-519-6
5	600.5	32.7	1053	5	US-09-724-519-2
6	456.5	24.9	338	5	US-09-723-153-4
7	456.5	24.9	780	5	US-09-723-153-2
8	94	5.1	311	5	US-09-739-449-10490
9	93.5	5.1	464	5	US-09-739-449-9635
10	92.5	5.0	1118	6	US-60-248-505-785
11	89	4.9	619	5	US-09-739-449-10081
12	86.5	4.7	636	5	US-09-739-449-9049
13	85	4.6	1008	5	US-09-308-453-2
14	84.5	4.6	605	5	US-09-739-449-9838
15	84.5	4.6	1312	1	PCT-US01-08117-95
16	83	4.5	1731	5	US-09-739-449-8331
17	82	4.5	3078	6	US-60-248-505-695
18	81	4.4	401	5	US-09-739-449-11084
19	80	4.4	317	5	US-09-739-449-12873
20	80	4.4	471	5	US-09-739-449-9886
21	77.5	4.2	521	5	US-09-739-449-979
22	77	4.2	926	5	US-09-421-124-187
23	77	4.2	926	5	US-09-421-124-187
24	76.5	4.2	277	5	US-09-755-456-7
25	75.5	4.1	317	5	US-09-739-449-11446
26	75.5	4.1	492	5	US-09-739-449-11322
27	75	4.1	414	5	US-09-739-449-10769

ALIGNMENTS

RESULT 1  
US-09-724-519-10  
; Sequence 10, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finer, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Mak, John  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: proliferation states  
; FILE REFERENCE: 1014A  
; CURRENT APPLICATION NUMBER: US/09724,519  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 09/592,037  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR APPLICATION NUMBER: 09/428,156  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-519-10

Query Match 33.4%; Score 612; DB 5; Length 368;  
Best Local Similarity 40.6%; Pred No. 1, 2e-46;  
Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;  
QY 4 GGNIVVVRFPNAREIDRGAQIVRMESNOTILTPPGAEKARKSGKTIHQPKAFA 63  
DB 16 GKNIVVVRFPNAREIDRGAQIVRMESNOTILTPPGAEKARKSGKTIHQPKAFA 67  
QY 64 FDRSYWFKNAPNAREIDRGAQIVRMESNOTILTPPGAEKARKSGKTIHQPKAFA 120  
DB 68 FDMVEGASTK-----QIDVTRVVCPLDVEINGVNGTCTIFAYGTGTGKFTTMEERS 120  
QY 121 ---YGKEH---GVPIRCODMFRFRIEQLQKKNLCTVSEVSYLETINVRDLNPST 172  
DB 121 PNEETWEEDPLAGIITLPHUKEKLT---DNGTEFSKVSLEYINEELFDLNPSS 176  
QY 173 --KGNLKVREUP--STGPPYVDLAKLVRSFOIENLMOEGNKAQVVAATNNETSSRH 228  
DB 177 DIVSELOMDFDRNKGVLIGLEIETVHNKDEYVLEGNKAQRTATLNNAYSSRH 236  
QY 229 AVPTLTQKHDEETKMDTE---KVAKISYDLGASRATSGATGAKGAEINRSL 285

Sequence 9003, Ap  
Sequence 9683, Ap  
Sequence 9539, Ap  
Sequence 10719, A  
Sequence 10719, A  
Sequence 1279, Ap  
Sequence 11271, A  
Sequence 28, Appl  
Sequence 189, App  
Sequence 189, App  
Sequence 1172, Ap  
Sequence 85, Appl  
Sequence 1262, A  
Sequence 10147, A  
Sequence 10147, A  
Sequence 11466, A  
Sequence 10546, A

Db 237 SVFSVTI-----HMKETIDGELVKIKGLNVLVLAGSENIGSGAVDKRAREAGNIQS 292  
QY 286 STLGRVITAAADSSCKKQKQLVPPYRDSVLTWLLKDSLGSGNSMTAMIAISPADINFE 345  
Db 293 LTLGRVITAVE-----RTPHPVYRSKLTIRLODSLGRTRTSIIATISPAISLNEE 345  
QY 346 TLSTLYADSAK 357  
Db 346 TLSTLEYAHRAK 357

## RESULT 2

US-09-724-519-4  
; Sequence 4, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finner, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Mak, John  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: proliferation states  
; FILE REFERENCE: 1014A  
; CURRENT APPLICATION NUMBER: US/09/724,519  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-519-4

Query Match 33.4%; Score 612; DB 5; Length 382;  
Best Local Similarity 40.6%; Pred. No. 1.8e-46;  
Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;

QY 4 GGNKVVVVRPPFNAREIDRGAKCIVRMEGNQITLTPPGAEEKARKSGKTMGPKAFA 63  
Db 16 GNIQVVVRCPPNLAERKSAHSIVECD-----PVRKEVSVRTGGGLAKSSRKYT 67  
QY 64 FDRSYWSFKNAPYARQEDLFQDLGVPLLDNAFYGNCFAYCGTGSGKSYSMG--- 120  
Db 68 FDMVFGASTK-----QIDVYRSVVCPILDVIMGYNCTIFAYCGTGCTKFTTMEGERS 120  
QY 121 ----YKHEH---GVIPRICODMFRINELOKQNKITCTVEVSLEYIENYRVDLLNPST 172  
Db 121 PNEYTWEEEDPLAGIIPRLHQIFELKLT---DNQTEFSVKVSLLEYINEELFDLLNPSS 176  
QY 173 --KGNLVAREHP--STGPYVEDLAKLVRSFQEIENLMDGECNKARTVAATNNMETSRS 228  
Db 177 DVSERLQWDFDPRNKRNGVYIIGLEEITVHNKDEYVGLLEGAARKRTAATLMAYSSRH 236  
QY 229 AVFTLTLOKWHDEETKMDTE--KVAKISLVLAGSERATSTGATGARLKEGAENRSL 285  
Db 237 SVFSVTI-----HMKETIDGELVKIKGLNVLVLAGSENIGSGAVDKRAREAGNIQS 292  
QY 286 STLGRVITAAADSSCKKQKQLVPPYRDSVLTWLLKDSLGSGNSMTAMIAISPADINFE 345  
Db 293 LTLGRVITAVE-----RTPHPVYRSKLTIRLODSLGRTRTSIIATISPAISLNEE 345  
QY 346 TLSTLYADSAK 357  
Db 346 TLSTLEYAHRAK 357

## RESULT 3

US-09-724-519-8

; Sequence 8, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finner, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Mak, John  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: proliferation states  
; FILE REFERENCE: 1014A

US-09-724-519-6  
; Sequence 6, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finner, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Mak, John  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: proliferation states  
; FILE REFERENCE: 1014A  
; CURRENT APPLICATION NUMBER: US/09/724,519  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR FILING DATE: 2000-06-12  
; PRIOR FILING DATE: 1999-10-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Human  
US-09-724-519-6

Query Match 33.4%; Score 612; DB 5; Length 513;  
Best Local Similarity 40.6%; Pred. No. 2.7e-46;  
Matches 151; Conservative 56; Mismatches 117; Indels 48; Gaps 10;

QY 4 GGNKVVVVRPPFNAREIDRGAKCIVRMEGNQITLTPPGAEEKARKSGKTMGPKAFA 63  
Db 16 GNIQVVVRCPPNLAERKSAHSIVECD-----PVRKEVSVRTGGGLAKSSRKYT 67  
QY 64 FDRSYWSFKNAPYARQEDLFQDLGVPLLDNAFYGNCFAYCGTGSGKSYSMG--- 120  
Db 68 FDMVFGASTK-----QIDVYRSVVCPILDVIMGYNCTIFAYCGTGCTKFTTMEGERS 120  
QY 121 ----YKHEH---GVIPRICODMFRINELOKQNKITCTVEVSLEYIENYRVDLLNPST 172  
Db 121 PNEYTWEEEDPLAGIIPRLHQIFELKLT---DNQTEFSVKVSLLEYINEELFDLLNPSS 176  
QY 173 --KGNLVAREHP--STGPYVEDLAKLVRSFQEIENLMDGECNKARTVAATNNMETSRS 228  
Db 177 DVSERLQWDFDPRNKRNGVYIIGLEEITVHNKDEYVGLLEGAARKRTAATLMAYSSRH 236  
QY 229 AVFTLTLOKWHDEETKMDTE--KVAKISLVLAGSERATSTGATGARLKEGAENRSL 285  
Db 237 SVFSVTI-----HMKETIDGELVKIKGLNVLVLAGSENIGSGAVDKRAREAGNIQS 292  
QY 286 STLGRVITAAADSSCKKQKQLVPPYRDSVLTWLLKDSLGSGNSMTAMIAISPADINFE 345  
Db 293 LTLGRVITAVE-----RTPHPVYRSKLTIRLODSLGRTRTSIIATISPAISLNEE 345  
QY 346 TLSTLYADSAK 357  
Db 346 TLSTLEYAHRAK 357

## RESULT 4

US-09-724-519-8  
; Sequence 8, Application US/09724519  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Finner, Jeffrey  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Mak, John  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Methods of screening for modulators of  
; TITLE OF INVENTION: cell proliferation and methods of diagnosing cell  
; TITLE OF INVENTION: proliferation states  
; FILE REFERENCE: 1014A

```

US-09-724-519-2

Query Match          32.7%  Score 600.5;  DB 5;  Length 1053;
Best Local Similarity 40.6%  Pred. NO. 7.6e-45;
Matches 151;  Conservative 55;  Mismatches 117;  Indels 49;  Gaps 11;

QY  4  GGNKVVYRVFPNAREIDRGACIVRMGNGQTIITPPGAEAKKSGKTIIMDGKAF 63
      11:||||| |||| | | | | | | | | | | | | | | | | | | | | | |
DB  16  GKNVYWRKFNFLNARKASNIHVED-----PYRVEYVETGLGKSRITII 67
      11:||||| |||| | | | | | | | | | | | | | | | | | | | | | |
QY  64  FDRSWSKDKMNPVAREQDFDLQGLLDNAPKGVNCFNCFAYGQTSCKSVNMG--- 120
      11:||||| |||| | | | | | | | | | | | | | | | | | | | | | |

```

[illegible]

```

; PRIOR APPLICATION NUMBER: 09/034,957
; FILING DATE: 2000-08-06
; NUMBER SEQ NOS.: 4
; SOFTWARE: FAST-SEQ for Windows Version 4.0
; SEQ ID NO. 4
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Human
; US-09-723-153-4

Query Match          24.9%; Score 456.5; DB 5; Length 338;
Best Local Similarity 38.5%; Pred. No. 5e-33;
Matches 115; Conservative 46; Mismatches 113; Indels 23; Gaps

OY 69 WSPDKNAPYARQEDL-FQDLGVPLLDNFAKGNICIFYAGTGSKSYNMGV---KE 124
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 50 WSEKLDVLDNLSQALVETAKQVQSDQAGTQCTYCTGCTGCTGCTGCTGCTGCT 109
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 125 HGVTRICQDMFRINELQNDKALCTVSYNIEVQVLLN-----PSTGNLKV 178
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 110 HGLIPRALQVFRNIEERPTH---AIVRVSYLEIYNESIFDLILSLPYVGSVT-PMTI 165
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 179 RHPSTGPVEDYDLAKLVRSFOETENLMDGNKATVAATNMNTRSSRSRAVFTLTOK 238
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 166 VEMPO-GVPIKGLSVHLTQSEEDAFSLFEGENRIIASHGMKNNSSRSCTFTLV--E 222
      :|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```
QY 239 WIDEETKMDTEKAKISLVLDLAGSERATSGATGARKKEGATINBSLTGSRVIAALADM 298
Db 223 AHSTLTSEKVIITSKINLVLDLAGSERLKGSGSQGVKLKEATYINKSLFEGAIIALGD- 281
QY 299 SSGQKKKNOLVPRDVSIVLTLLKDSLGNSWMTAMTAATSPADINFEETSLTLRYADSAAK 357
Db 282 -----QKRDRHIPFROCKLTHALKDSLGNCNMVLTNYIYGERAAQLEETLSLRSASRMK 335

RESULT 7
US-09-723-153-2
; Sequence 2, Application US/09723153
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1048
; CURRENT APPLICATION NUMBER: US/09/723,153
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 09/634,957
; PRIOR FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Human
US-09-723-153-2
```

```
Query Match 24.98; Score 456.5; DB 5; Length 790;
Best Local Similarity 38.5%; Pred. No. 1.7e-32;
Matches 115; Conservative 48; Mismatches 113; Indels 23; Gaps 8;

QY 69 MSFQKNAPYARQEDL-FQDLGVPLLDNAFKGNYNCPAYGQTGSGKSYMMYG--RE 124
Db 53 MSFGLDVLHDAQSOLIVETVAKDVYVSQALDYGNTIMCYGTGAGKTYTMGATENYKH 112
QY 125 HGVIPTICQDMFRINELQDKNLACTVEVSYLEYINERVRLDN-----PSTKGNLKV 178
Db 113 RGILRALQGVFRMEIERPTH---ATVRSYLEIYNESLFDLLSLTPYGPVSIVT-PMTI 168
QY 179 REHPSTGPVVEDLAKLVRSQEIENLMDGNKARTVAATNNNETSSRSRHVFTLTUTOK 238
Db 169 VENPQ-GVFIKGLSVHLTSGEEDARSLFEGEYRRIATSMNKNSSRSCHIFTIYL--E 225
QY 239 WIDEETKMDTEKAKISLVLDLAGSERATSGATGARKKEGAEINBSLTGSRVIAALADM 298
Db 226 AHSTLTSEKVIITSKINLVLDLAGSERLKGSGSQGVKLKEATYINKSLFEGAIIALGD- 284
QY 299 SSGQKKKNOLVPRDVSIVLTLLKDSLGNSWMTAMTAATSPADINFEETSLTLRYADSAAK 357
Db 285 -----QKRDRHIPFROCKLTHALKDSLGNCNMVLTNYIYGERAAQLEETLSLRSASRMK 338
```

```
RESULT 8
US-09-730-449-10490
; Sequence 10490, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 10490
; LENGTH: 331
; TYPE: PRT
```

```
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10490

Query Match 5.18; Score 94; DB 5; Length 331;
Best Local Similarity 19.48; Pred. No. 0.16;
Matches 77; Conservative 48; Mismatches 107; Indels 164; Gaps 20;

QY 47 KARKSKGTIMQPKAFADRSYV-SFDKNAPYARQE-----DLFQDLGVP- 91
Db 2 KYNHCKT---GLKLPAISLGLMHNFNDTPHQTQAICRCRAFDLGIHFDLANNYPGP 58
QY 92 -----LLDNAFGYNNCIF-----AYGQTGSGKSYMMYGKGVIPR 130
Db 59 GSAETAFGEILKTDPRGYDEMIISSKAGYNMWPGPYGEWSRK-----YLIS 106
QY 131 IQQDMFRINELQDKNLCTVEVSYLEYINERVRLDNLNPSTKGNLKVREHPSTGTYVE 189
Db 107 SCDSQSKRMG-----LDYVDIFYSHR-----FDPNT-----PLEE 136
QY 190 DLAKL--VYRSQEIENLMDGNKARTVAATNNNETSSRSRHV-----TLT 234
Db 137 TCGALQGVIRS-----GKALYGVIGISSYNSKSRTRREAAAILKDLGTPCIIHQPSYS 185
QY 235 LTQKWHDEETKMDTEK---VAKISLVDLA-----GSRATSTGATCARLKE 277
Db 186 MINHWIETDGLVDTLEELGIGSVFSLAQGLMTTKYLGVPDGSRASQSKSLNPAFLNE 245
QY 278 -GAETNRSLSLT---GRVIAALADMSSGQKKKNOLVPRDVSIVLTLLKDSLGNSMTAM 332
Db 246 RNVENIRALNSIAERRGOTLAQMA-----IAWVLR---GGRITSAL 283
QY 333 IA-----AISPADINFEETSLTLRYADSA 356
Db 284 IGASRVEQVEDCVKALDNAEFSTEELAEIDRYAKDA 319

RESULT 9
US-09-739-449-9635
; Sequence 9635, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9635
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9635

Query Match 5.18; Score 93.5; DB 5; Length 464;
Best Local Similarity 20.28; Pred. No. 0.29; Mismatches 113; Indels 113; Gaps 16;
Matches 72; Conservative 58;

QY 7 IKVVRVFRNAREIDRGAKCIVYRMGNOTILTPPPGAEEKARKSGK-----53
Db 81 LEVESKRTIISRNVLAVDVELKLDQDPEVFPPLARKAMFSPDGDGNRVGALIS 140
QY 54 -----TIMQPKAFADRSYVSDKNAPYARQEDLFDLGVPLLDNAFKGNYNCPAYG 108
Db 141 LGDRVTRAEADPRSVFTVLSVWIND-----AEKSAVSAKVAIFESLFI---184
QY 109 OTGSGKSYMMYGKGVIPRITICODMFRINELQDKNLCTVEVSYLEYINERVRLD 168
Db 185 QTNDSSS-----QRVVDLAKRIEDAR--ENVI-----IAEKVADFR 220
```

```

: TYPE: PRT
: ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10801

Query Match          4.9%; Score 89; DB 5; Length 619;
Best Local Similarity 22.3%; Pred. No. 1.1;
Matches 75; Conservative 45; Mismatches 133; Indels 84; Gaps 16;

OY 78 YAQ-EDFDGLGVPLDNLKAFGYNCTFAYGOTGSKSY---SNMGVY-KHGVYIPRI 131
DB 48 YQYVSLDSRGFTGTLLTFEQAKRGFDGQVFSQALAAALDELAKNLGLSENLAWLI 107
OY 132 COD-MFRRIENELAKKNICTTVEYSVLEINRVERDLLNPSTKGNLKVREHPSPGYVED 190
DB 108 AEDPAFYSVNG-QFDRLN-----FSERLR-----NSGFRE---DDYIKE 142
OY 191 LAKLVVRSFOEINLMDGSKARTVAATNNMETSRSRSHAVTL-----T 234
DB 143 RSKVAVRS-QIVAVSDGPAAPVLDAVKQVRNQRVADYVILSNVIPPVKAGQDV 200
OY 235 LTKQWDEETKMKTEKYAKISVLDAGSEKATSTGATGAKLGKAEINR-SLSTLGR--- 290
DB 260 LTFPWEETNKSRYNRFKFTFYVLLKPSDIAEPASVDTAQDIADYINSHKDSFRTAGRTV 260
OY 291 -----VIALADMSKQKKNLQVYR-----DSVLTWLLKSLGGNSWMTATIA 335
DB 261 EQLTTPDKEMAAAEQILGNWTVQVYVKKQGTASDVTLGSEFTKUTIPDQSIADAAFA 320
OY 336 I-----SP-ADINFEETL-----STLRVYDSAK 357
DB 321 IQKDGGVSPVWDGFCPIILLRVITGKIPKTRTLLDAK 357

```

```

RESULT 12
US-09-739-449-9049
: Sequence 9049, Application US/09739449
: GENERAL INFORMATION:
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTOR: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-10(15490)C
: CURRENT APPLICATION NUMBER: US/09/739,449
: CURRENT FILING DATE: 2000-12-19
: PRIOR APPLICATION NUMBER: US 09/514,000
: PRIOR FILING DATE: 2000-02-23
: NID/SEQ ID NO: 9049
: SEQ ID NO: 9049
: LENGTH: 638
: TYPE: PRT
: ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9049

Query Match 4.7%; Score 86.5; DB 5; Length 636;
Best Local Similarity 24.3%; Pred. No. 1:9;
Matches 82; Conservative 40; Mismatches 153; Indels 63; Gaps 18;

QY 17 NAREIDGAKICIVMEGNQITLTPPGAEEK--AR-KSKGTMD-----GPRAFAP----- 64
DB 162 NAGALD--ARLIVLNDNDMSITAPTGTGAMGAYLARLASCRTGYGFRDFGKTIATYIGKTI 219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 DRSYWSFDKNAPWAEQEDLFQDUG-----VPLDNKAFGYNCNCFAYGQ 109
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 220 DRAITRATHANGVTGGLTIFELUGYHGPIDIGSHDILPLFVLRWVDKNGKGPLIIRV 279
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 110 TGSQKSY--SMWGYCKEHSVTRICODMPPRINELQDNKILCTCTVEYSLEYLIERVEDL 167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 280 TOKGGKGYAPAAADKYHGV-----NKFQDVTGQAQKAKNPAP-----SYTSVFAPAA--L 327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 LNESTKGN--LKVREHSPGTVPVEDIAKLIV-VRSFQEIENLMDG---NKARTVAATNN 221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 328 IQEATLDEKIGIVTAAMPNGTCTGLQKMAELFSGRTF-----DVGIAEQHAYTFAA--GLA 379
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
;
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9838

Query Match
Best Local Similarity 20.4%; Pred. No. 2.6;
Matches 82; Conservative 59; Mismatches 165; Indels 95; Gaps 19;

QY 222 ETSRSHAVFTLTQKHDEETKMDTEKVAKISLVDLGASERATSTGATGARLKEAGEI 281
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 380 ADGKPFKALYSTFLQGYDQL-----VHDVAIQSLVFRFPIDRAGVGADGP-THAGSFD 434
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 282 NRSLSLTL-GRVIAALADSSGKKQKNQLVYPROSVLTW 318
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 435 TTFLATLPGMVMAAADEALKHNVRTAAAYDEGPISF 472
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-09-308-453-2
; Sequence 2, Application US/09308453
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GMBH
; TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum and
; FILE REFERENCE: BMID9924US
; CURRENT APPLICATION NUMBER: US/09/308,453
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Clostridium histolyticum
US-09-308-453-2

Query Match
Best Local Similarity 19.0%; Pred. No. 4.8;
Matches 63; Conservative 40; Mismatches 143; Indels 86; Gaps 13;

QY 53 KTMGDKAFAPDSRWSP-----DKNAPNVAQEDLFQDLGVPLLDNAFK-- 98
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 483 KTLNSG-----IDSDWMEYNGFVAHYLYEKDMPPTIRMKKALINDVKSYEIIKL 537
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 99 -----GYNCCI--FAYGOTSGSKSYSMGKGHEGVIPRICODMFRINELQKOKNU 148
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 538 SDDANKTEYQHHLVDYKQAGLPLVSDDLADHG--KKASEVISEIKSAASUTNT 595
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 149 TCTVENS-VLEIYNERVRLDNLSTKGNLKVREHSTGPYVEDLAKLWVSQBEIENLMD 207
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 596 SVTAERKSQYENTFLR-----GTYGETSGEFGEDWDEMSKKLID 634
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 208 -----EGNKARTVAATNMNETSSRS---HAVFTLTQKHDEETKMDTEKVAK 253
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 635 GTLESIAKNSGSKYKLTATFYATNRYVTSDNKQYDVVPHGVLTDNGDISNKAPIKATV- 693
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 ISLVDLASERATSTGATGARLKEAGEINRSLSLGRVIAALADSSG--KQKNQLVYK 311
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 694 -----GPSTGAVGRNIEFSGKSKDED--GKIVSYDMDGDDGATSGKNSHVAY 740
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 312 RDS---VLMTLLKDSLGNSMTAIAASPAD 340
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 741 KRAGTYNVLKVTDDKGATATESFTIEIKNED 772
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-739-449-9838
; Sequence 9838, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkley, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9838
; LENGTH: 605
```

```
;
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9838

Query Match
Best Local Similarity 20.4%; Pred. No. 2.6;
Matches 82; Conservative 59; Mismatches 165; Indels 95; Gaps 19;

QY 5 GNIKVVYVRFPFNAREIDRGAKCTVRMEG-----NOTILTPPPGAEEKARKSGKGTMDGPK 60
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 244 GSMAAAVAVFRSNAME-----RLRLGDAEQNRTL--SEQERNERERTAAKDAAD--I 292
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 AFAFDRSYNSFDKNAPNVAQEDLFQDLGVPL---LDNAFGYNCNCFPAYGOTSGSKSYS 117
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 OFAVD-----SLAKGLAHL--DCDLNRYIDTPTVTRIDRLRDNFNNSVAKLNAALSTVGQN 347
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 MMGCKBHGVPRIQODMFRINELQK-----KNLCTVETVSYLEIYNERVRLDNL 169
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 348 ALADAGAGEIRQASDLARRTEQQAASVEETAAALEEITTVTKDSARRA--EEVGRIVD 405
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 PSTGNLKVREHSTGPYVEDLAKL---VVSRSQBEIENL----- 206
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 406 -RARGNAE-----QSGVIVVEDAVRAMEGTEKSSSEISNIIIGVIDEIAFOTNLALNAGVE 459
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 207 ----DEGNKARTVAATNMNETSSRS---HAVFTLTQKHDEETKMDTEKVAKISLV 257
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 460 AARAGEAGKGFVAVQAEVRELAORSANAAKAIRT-----INASTSOVSQSGVELV 509
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 DLAGSERATSTGATGARLKEAGEINRSLSLGRVIAALADSSGKQKNQLVYPRD--- 313
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 510 GNAAKALETI-----VREVEQIEINHIDA---IVTSSRQSTGLQETINTINTIDGQTQ 559
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 314 SVLTMLLKDSLGNSMTAIAASPADINFEETLSLRYAD 354
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 560 QNAAVEEQTAASHGLASPAALNELLAQFOLAATAATRQAE 600
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
PCT-US01-08117-95
; Sequence 95, Application PC/TUS0108117
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: TUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LO, Dyoung Alina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: ACIZAL, Talda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MAIHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AD-TOUNG, Janice
; APPLICANT: REDDI, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/08117
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 95
; LENGTH: 1312
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 4880891CD1
PCT-US01-08117-95

Query Match
Best Local Similarity 20.4%; Pred. No. 2.6;
Matches 82; Conservative 59; Mismatches 165; Indels 95; Gaps 19;

QY 5 GNIKVVYVRFPFNAREIDRGAKCTVRMEG-----NOTILTPPPGAEEKARKSGKGTMDGPK 60
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 244 GSMAAAVAVFRSNAME-----RLRLGDAEQNRTL--SEQERNERERTAAKDAAD--I 292
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 AFAFDRSYNSFDKNAPNVAQEDLFQDLGVPL---LDNAFGYNCNCFPAYGOTSGSKSYS 117
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 OFAVD-----SLAKGLAHL--DCDLNRYIDTPTVTRIDRLRDNFNNSVAKLNAALSTVGQN 347
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 MMGCKBHGVPRIQODMFRINELQK-----KNLCTVETVSYLEIYNERVRLDNL 169
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 348 ALADAGAGEIRQASDLARRTEQQAASVEETAAALEEITTVTKDSARRA--EEVGRIVD 405
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 170 PSTGNLKVREHSTGPYVEDLAKL---VVSRSQBEIENL----- 206
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 406 -RARGNAE-----QSGVIVVEDAVRAMEGTEKSSSEISNIIIGVIDEIAFOTNLALNAGVE 459
; : : : : : : : : : ~~~~~~
QY 207 ----DEGNKARTVAATNMNETSSRS---HAVFTLTQKHDEETKMDTEKVAKISLV 257
; : : : : : : : : : ~~~~~~
DB 460 AARAGEAGKGFVAVQAEVRELAORSANAAKAIRT-----INASTSOVSQSGVELV 509
; : : : : : : : : : ~~~~~~
QY 258 DLAGSERATSTGATGARLKEAGEINRSLSLGRVIAALADSSGKQKNQLVYPRD--- 313
; : : : : : : : : : ~~~~~~
DB 510 GNAAKALETI-----VREVEQIEINHIDA---IVTSSRQSTGLQETINTINTIDGQTQ 559
; : : : : : : : : : ~~~~~~
QY 314 SVLTMLLKDSLGNSMTAIAASPADINFEETLSLRYAD 354
; : : : : : : : : : ~~~~~~
DB 560 QNAAVEEQTAASHGLASPAALNELLAQFOLAATAATRQAE 600
; : : : : : : : : : ~~~~~~
```



		Best Local Similarity    22.5%; Pre; Mismatch 70; Indels 83; Gaps 15; Matches 55; Conservative 36;
QY	60 KAFADRSVDFSKNAPFARQEDFDQLGVLIDNAKGN-NCIFAYQTGSKSYSM I18         :   :	
Db	342 KLFLRVHLKGDF-NIESGAWKQVTVDGLPILNS-AGNVKC-----ATKK 388         :   :	
QY	119 MTGKE-----HGVPRICDPMFRINELQ-KKNLTCTVEVSLEYINE 162       :   :	
Db	389 YLIGFEVCSANIEFOMALPERVNKKOCAC-ENYKLKRENT--EIREKNESE 444         :   :	
QY	163 R-----VRDLUNPTKNGLKVRHPSTG-----PYVEDAKLV-VSRQEITE 203         :   :	
Db	445 RNIIIPREEKPIDEI--ERKENIK---PLSGSKNNLLSIPTHSDQEVNIKKPEDNE 498         :   :	
QY	204 NMD-----EGNKARTYAATNNINTSSISHAFTVTTUKKHDEETK 245         :   :	
Db	499 NLDDMKDDTTRVDESLSMIKYAEKAASGDNETKDDEDEEA-----ESEEEEEESE 552         :   :	
QY	246 NOTE 249 :::	
Db	553 EDED 556 :::	

Search completed: April 25, 2001, 10:16:32  
Job time: 303 sec

**THIS PAGE BLANK (USPTO)**



```

; LOCATION: (443)...(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (602)...(784)
; OTHER INFORMATION: tail domain
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (713)
; OTHER INFORMATION: polymorphic variant #1 Val -> Ile
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (762)
; OTHER INFORMATION: polymorphic variant #2 Asp -> Glu
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (774)
; OTHER INFORMATION: polymorphic variant #3 Glu -> Asp
; US-09-235-416-1
; PCT-US99-01355-1

```

```

Query Match      100.0%; Score 938; DB 1; Length 784;
Best Local Similarity 100.0%; Pred. No. 5.3e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QEOSLLRHSVNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPHFRGKDSDMFYARRE 60
    |||||
Db 602 QEOSLLRHSVNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPHFRGKDSDMFYARRE 661
    |||||
Qy 61 AASAILGLDQKISHLTDDLDALFDVQKARVRRLGVEDNEDSDSSQSFVPVRDKYMSNG 120
    |||||
Db 662 AASAILGLDQKISHLTDDLDALFDVQKARVRRLGVEDNEDSDSSQSFVPVRDKYMSNG 721
    |||||
Qy 121 TIDNFSLOTALTMPGTPRSDDDGDALFFGDKSKODASNVVDELRQQAQMEELKTKAK 180
    |||||
Db 722 TIDNFSLOTALTMPGTPRSDDDGDALFFGDKSKODASNVVDELRQQAQMEELKTKAK 781
    |||||
Qy 181 QEF 183
    |||
Db 782 QEF 784

```

```

RESULT 2
US-09-235-416-1
; Sequence 1, Application US/09235416A
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 18557C-0007100S
; CURRENT APPLICATION NUMBER: US/09/235,416A
; CURRENT FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: WO PCT/US99/01355
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 60/072,361
; EARLIER FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)...(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; NAME/KEY: DOMAIN

```

```

; LOCATION: (358)...(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (443)...(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (602)...(784)
; OTHER INFORMATION: tail domain
; US-09-235-416-1

```

```

Query Match      100.0%; Score 938; DB 16; Length 784;
Best Local Similarity 100.0%; Pred. No. 5.3e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QEOSLLRHSVNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPHFRGKDSDMFYARRE 60
    |||||
Db 602 QEOSLLRHSVNSQLGSPAPGRHRTLSKAGSDADGDSRSDSPLPHFRGKDSDMFYARRE 661
    |||||
Qy 61 AASAILGLDQKISHLTDDLDALFDVQKARVRRLGVEDNEDSDSSQSFVPVRDKYMSNG 120
    |||||
Db 662 AASAILGLDQKISHLTDDLDALFDVQKARVRRLGVEDNEDSDSSQSFVPVRDKYMSNG 721
    |||||
Qy 121 TIDNFSLOTALTMPGTPRSDDDGDALFFGDKSKODASNVVDELRQQAQMEELKTKAK 180
    |||||
Db 722 TIDNFSLOTALTMPGTPRSDDDGDALFFGDKSKODASNVVDELRQQAQMEELKTKAK 781
    |||||
Qy 181 QEF 183
    |||
Db 782 QEF 784

```

```

RESULT 3
US-09-654-850-1
; Sequence 1, Application US/09654850
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Goldstein, Lawrence S. B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Identification and Expression of a Novel Kinesin Motor
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 18557C-0007100S
; CURRENT APPLICATION NUMBER: US/09/654,850
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/235,416
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 60/072,361
; PRIOR FILING DATE: 1998-01-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Thermomyces lanuginosus
; FEATURE:
; OTHER INFORMATION: TL-gamma ATP-dependent plus end-directed
; OTHER INFORMATION: microtubule motor protein
; NAME/KEY: DOMAIN
; LOCATION: (1)...(357)
; OTHER INFORMATION: kinesin-like microtubule motor domain
; NAME/KEY: DOMAIN
; LOCATION: (358)...(442)
; OTHER INFORMATION: neck domain links motor domain to stalk domain
; NAME/KEY: DOMAIN
; LOCATION: (443)...(601)
; OTHER INFORMATION: stalk domain, unc-104 family domain
; NAME/KEY: DOMAIN
; LOCATION: (602)...(784)
; OTHER INFORMATION: tail domain
; US-09-654-850-1

```

```
Query Match          100.0%; Score 938; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 5.3e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QEOSLLRHSYNSGSPARCDRTLSKAGSDAGDSRSDPLPHFRGKSDSMFYARRE 60
    |||
Db 602 QEOSLLRHSYNSGSPARCDRTLSKAGSDAGDSRSDPLPHFRGKSDSMFYARRE 661
    |||

Qy 61 AASATLGLDQKISHLTDELDAFDVQKARVRGLVEDNEDSDSQSPFVRDKYMSNG 120
    |||
Db 662 AASATLGLDQKISHLTDELDAFDVQKARVRGLVEDNEDSDSQSPFVRDKYMSNG 721
    |||

Qy 121 TIDNESLDATMTMCTPRSDGDAALFFGDKKSKODASNVNVEELRQQAOMEALAKTAK 180
    |||
Db 722 TIDNESLDATMTMCTPRSDGDAALFFGDKKSKODASNVNVEELRQQAOMEALAKTAK 781
    |||

Qy 181 QEF 183
    |||
Db 782 QEF 784
    |||

RESULT 4
US-09-417-507-42125
: Sequence 42125, Application US/09417507
: GENERAL INFORMATION:
: APPLICANT: KEITH G. WEINSTOCK ET AL.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
: FILE REFERENCE: 1999-10-14
: CURRENT FILING DATE: 1999-10-14
: CURRENT APPLICATION NUMBER: US/09/417,507
: NUMBER OF SEQ ID NOS: 44312
: SEQ ID NO 42125
: LENGTH: 1000
: TYPE: PRT
: ORGANISM: A. fumigatus
US-09-417-507-42125

Query Match          50.4%; Score 472.5; DB 18; Length 1000;
Best Local Similarity 57.4%; Pred. No. 1e-38;
Matches 101; Conservative 22; Mismatches 28; Indels 25; Gaps 4;

Qy 32 SDAXGD-SRSDSPLPFRGKSDSMFYARREASAILGLDQKISHLTDELDAFDVQKA 90
    |||
Db 1 SELDQGSRSNSPQAQAGDAEFMYARREAVSAILGDH-ISMHPDELDALFDVQKY 59
    |||

Qy 91 RVRGLVEDNEDSDSQSPFVRDKYMSNGTIDNESLDATMTMCTPRSDGDAALFFGDK 150
    |||
Db 60 RATRGLVEDNEDSDLSLSPFVRDKYMSNGTIDNESLDATMTMCTPGQYDSEQNGSD 119
    |||

Qy 151 ---KKSQD---
    |||
Db 120 FTLLQAARDNQRLDKQKEFKKRLRIAEASDQADADELRLEKRMEEALSTKEEY 175
    |||

RESULT 5
US-09-417-507-33143
: Sequence 33143, Application US/09417507
: GENERAL INFORMATION:
: APPLICANT: KEITH G. WEINSTOCK ET AL.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
: FILE REFERENCE: 1999-10-14
: CURRENT FILING DATE: 1999-10-14
: CURRENT APPLICATION NUMBER: US/09/417,507
: NUMBER OF SEQ ID NOS: 44312
: SEQ ID NO 33143
: LENGTH: 466
: TYPE: PRT
: ORGANISM: A. fumigatus
: FEATURE:
```

```
: NAME/KEY: UNSURE
: LOCATION: (412)
: OTHER INFORMATION: Identity of amino acid sequences at the above locations are un
US-09-417-507-33143

Query Match          10.9%; Score 102; DB 18; Length 466;
Best Local Similarity 23.3%; Pred. No. 0.11;
Matches 45; Conservative 35; Mismatches 75; Indels 38; Gaps 8;

Qy 24 DRTLSKAG-SDAGDSRSDSPLPFRGKSDSMFYARREASAILGLD----QKISHLTDD 78
    |||
Db 168 DLRLHLPDQGFQIDSTIRAPLPARKARDAMVVENAEAPQVLDVPTADRDLLATVAPE 227
    |||

Qy 79 ELDAHLDPDQKARVRGLVEDNED--SDSQSPFVRDKYMSNGT-----IDNES-- 126
    |||
Db 228 ETEMEEDDISIAETKRGVLLDHHYFSDSDSHLPARPKRLPKGTSEYQSAWFIIDVSDS 287
    |||

Qy 127 -----LDATMTMCTPRSDGDAALFFGDKK-----SKODASNVNVEELRQQA 169
    |||
Db 288 GSDTEEEEDQDEAMAM-DTAGNPDG--VFPDRQDAMTEAGSPSEYPOSEMFLDPSPEDE 343
    |||

Qy 170 AQMEELAKTAKOE 182
    |||
Db 344 AQOLEEYRASRRK 356
    |||

RESULT 6
US-09-733-089-17684
: Sequence 17684, Application US/09733089
: GENERAL INFORMATION:
: APPLICANT: Deacon, Stanton B.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Lutfiyeva, Linda L.
: APPLICANT: McIninch, James
: APPLICANT: Wu, Wei
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: FILE REFERENCE: 38-21(15300)D
: CURRENT APPLICATION NUMBER: US/09/733,089
: CURRENT FILING DATE: 2000-12-11
: PRIOR APPLICATION NUMBER: US 09/474,435
: PRIOR FILING DATE: 1999-12-28
: PRIOR APPLICATION NUMBER: US 09/654,617
: PRIOR FILING DATE: 2000-09-05
: PRIOR APPLICATION NUMBER: US 09/620,392
: NUMBER OF SEQ ID NOS: 24143
: SEQ ID NO 17684
: LENGTH: 845
: TYPE: PRT
: ORGANISM: Oryza sativa
US-09-733-089-17684

Query Match          10.6%; Score 99; DB 21; Length 845;
Best Local Similarity 23.7%; Pred. No. 0.54;
Matches 46; Conservative 32; Mismatches 66; Indels 50; Gaps 8;

Qy 21 GHORTLSKAGSDAGDSRSDPLPHFRGKSDSMFYARREASAILGLDQKISHLTDD 80
    |||
Db 519 GEIDHSTOK---EASMDSDSKRRKRRESGHDSS---SRVQSSSLRNRLCHQLJLQEC 572
    |||

Qy 81 DAL-----FDVQKARV-----RGLVEDNEDSDSQ 107
    |||
Db 573 DDLKYSSTNDFKATSMKRLFTISVQLQVPTLPYPASPLASSETNRLVQGRNSCR 632
    |||

Qy 108 SSSFV---RDKYMSNGTIDNESLDATMTMCTPRSDGDAAL-FFGDKK---SKODASNV 160
    |||
Db 633 NIILDSUNDEDTYFANVONIGANTTVLV---USDGDSVASFVDEKSDSKQANYYI 688
    |||

Qy 161 DVEELRQQAQME 174
    |||
```

```

;      :  | : | | : |
Db 689 EESVLPQHAQQOE 702

RESULT 7
US-08-302-166-4
; Sequence 4, Application US/08302166
; GENERAL INFORMATION:
; APPLICANT: PRENDERGAST, Kenneth Francis
; TITLE OF INVENTION: IMAGING COMPOSITION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTIE, PARKER & HALE
; STREET: P. O. BOX 7068
; CITY: PASADENA
; STATE: CA
; COUNTRY: USA
; ZIP: 91109-7068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,166
; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 14481.5
; FILING DATE: 08-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 93 01221.9
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 93 13965.7
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CARNER, HAYDEN A.
; REFERENCE/DOCKET NUMBER: 22653
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 818-795-9900
; TELEFAX: 818-577-8800
; TELEX: ITT 4995129
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1030 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: YES
US-08-302-166-4

Query Match 10.5%; Score 98.5; DB 7; Length 1030;
Best Local Similarity 24.3%; Pred. No. 0.81;
Matches 45; Conservative 32; Mismatches 59; Indels 49; Gaps 9;

Qy 6 LRHSVNSQLSGSPAPGRHRTLSKAGSDADGDSRSDSLPHFRKDSDFMYARREASAI 65
Db 693 LRHSKDNDSGDGPA-----ESMANPDSNSKGTG-----KQDNDAKATKDSNS 739
Qy 66 LGLDQKISHLTDELDALFDVQKARAVRGLVEDNEDSDSQSFPPVKYMSN----GT 121
Db 740 DG-----TSSATGDTTDAV-----DREINKVPED-----RUKTVGSKDGGGE 777
Qy 122 IDN-FSLDTAITMPTGTPRSDDDGALFFGDKKSKQASNVYVELRQOQAAEALK 177
Db 778 VDSANKDAAVTVGCDREINSAG-----GSTDNRSKNDTEKNGASTPDSQSEDA 831
Qy 178 TAFEQ 182
Db 832 LSKTE 836

;      :  | : | | : |
Db 689 EESVLPQHAQQOE 702

RESULT 8
US-60-173-464-5644
; Sequence 5644, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5644
; LENGTH: 1735
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-5644

Query Match 10.5%; Score 98.5; DB 23; Length 1735;
Best Local Similarity 23.1%; Pred. No. 1.7;
Matches 42; Conservative 24; Mismatches 73; Indels 43; Gaps 6;

Qy 8 HSVTNSQLSGSPAPGRHRTLSKAGSDADGDSRS-----DSPLPFRKDSDFMYARRE 60
Db 29 HSGSGSGSGS-----SGSDSDSSSGNSGDSGRSPEDKSLSVAGFPPTAA 76
Qy 61 AASAILGLDQKISHLTDELDALFDVQKARAVRGLVEDNEDSDSQSFPPVKYMSNG 120
Db 77 AAQA-----DSKNGFTDDOESSDG-----SSGSDSDSDADGSPDQR---NQ 117
Qy 121 TIDNFSLDITMPTGTPRSDDDGALFFGDKKSKQASNVYVELRQOQAAEALKTAK 180
Db 118 SINNAWTSSTSLPRKPEQNEEDNET-----EAGQQQPASDASDESSDSSANVSPTSSSS 172
Qy 181 QE 182
Db 173 SE 174

;      :  | : | | : |
Db 689 EESVLPQHAQQOE 702

RESULT 9
US-60-167-217-7043
; Sequence 7043, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; FILE REFERENCE: CL000155
; CURRENT APPLICATION NUMBER: US/60/167,217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7043
; LENGTH: 1736
; TYPE: PRT
; ORGANISM: Drosophila
US-60-167-217-7043

Query Match 10.5%; Score 98.5; DB 23; Length 1736;
Best Local Similarity 23.1%; Pred. No. 1.7;
Matches 42; Conservative 24; Mismatches 73; Indels 43; Gaps 6;

Qy 8 HSVTNSQLSGSPAPGRHRTLSKAGSDADGDSRS-----DSPLPFRKDSDFMYARRE 60
Db 30 HSGSGSGSGS-----SGSDSDSSSGNSGDSGRSPEDKSLSVAGFPPTAA 77
Qy 61 AASAILGLDQKISHLTDELDALFDVQKARAVRGLVEDNEDSDSQSFPPVKYMSNG 120
```

```

Db      78 AAQA-----DSKNTGFTDQDESSSDG-----SSGSDDSDAEGPSDQR--NQ 117
QY      121 TIDNFSLDATIMPTGRSDGDGDLFGDKKSKODASNVDELRLQQOAAEALKTKAK 180
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:
Db      119 SINNAHTSSLPKPEONEEDNET-----BAGQOQPASDADESSESSANVSPTSSESS 173
QY      181 QE 182
       |
Db      174 SE 175

RESULT 10
US-60-191-637-6927
; Sequence 6927, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venetier, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6927
; LENGTH 1883
; TYPE PRS
; ORGANISM: DROSOPHILA
US-60-191-637-6927

Query Match          10.5%   Score 98.5;   DB 23; Length 1883;
Best Local Similarity 23.1%; Pred. No. 2;
Matches 42; Conservative 24; Mismatches 73; Indels 43; Gaps 6;

QY      8 HSTVNSQLSGSPAPGRHDRTLSKAGSDADGDSRS-----DSPLPFRGKDSMFYARRE 60
Db      29 HSUGSGSGSGS-----SSGSDDSDSSGSPEDKSLSVAGFPPTAA 76
QY      61 AASAILGLQOKISHLTDELDALFDVOKARVGLVDNEGDSOSGSSPPRYKYMSNG 120
Db      77 AAQA-----DSKNTGFTDQDESSSDG-----SSGSDDSDAEGPSDQR--NQ 117
QY      121 TIDNFSLDATIMPTGRSDGDGDLFGDKKSKODASNVDELRLQQOAAEALKTKAK 180
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:
Db      118 SINNAHTSSLPKPEONEEDNET-----BAGQOQPASDADESSESSANVSPTSSESS 172
QY      181 QE 182
       |
Db      173 SE 174

RESULT 11
US-60-191-681-5421
; Sequence 5421, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLETIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5421
; LENGTH 1883
; TYPE PRT
; ORGANISM: DROSOPHILA
US-60-191-681-5421

Query Match          10.5%   Score 98.5;   DB 23; Length 1883;

```





---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:16:32 ; Search time 7.94 Seconds  
(Without alignment)  
77.702 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_602\_784

Perfect score: 938

Sequence: 1 QEQSLRHVTSNGLSPAP.....ELRQQQAEEALKTAQEF 183

Scoring Table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 17899 seqs, 3154390 residues

Total number of hits satisfying chosen parameters: 17899

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76.5	8.2	1503	5	US-09-649-996-14
2	73.5	7.8	759	6	US-09-792-024-76
3	73.5	7.8	393	6	US-60-248-505-1134
4	72.5	7.7	425	6	US-60-248-505-737
5	69.7	7.4	1758	5	US-09-739-449-11627
6	68.5	7.3	694	5	US-09-739-449-9792
7	68.7	7.2	492	5	US-09-739-449-10209
8	68.7	7.2	693	5	US-09-336-910A-3
9	68.7	7.2	714	5	US-09-792-024-117
10	67.7	7.1	1042	5	US-09-792-024-106
11	66.5	7.1	529	5	US-09-739-449-8477
12	66.5	7.1	806	5	US-09-739-449-11845
13	66.5	7.0	884	5	US-09-739-449-12024
14	65.5	7.0	814	5	US-09-832-248-3492
15	65.5	7.0	727	5	US-09-832-246-2
16	65.5	7.0	727	5	US-09-832-246-2
17	64.5	6.9	422	5	US-09-781-417-85
18	64.5	6.9	395	5	US-09-739-449-11881
19	64.5	6.9	648	5	US-09-813-408-26
20	64.5	6.9	695	5	US-09-806-194-10
21	64.5	6.9	695	5	US-09-806-194-12
22	64.5	6.9	695	5	US-09-806-194-11
23	64.5	6.9	697	5	US-09-806-194-16
24	64.5	6.9	697	5	US-09-806-194-18
25	64.5	6.9	697	5	US-09-806-194-20
26	64.5	6.9	976	5	US-09-628-359-18
27	64	6.8	281	5	US-09-739-449-10656

Sequence 12619, A  
Sequence 12453, A  
Sequence 8567, Ap  
Sequence 12243, A  
Sequence 78, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 11, Appl  
Sequence 112, Ad  
Sequence 309, App  
Sequence 165, App  
Sequence 87, Appl  
Sequence 114, App  
Sequence 9137, Ap  
Sequence 77, Appl  
Sequence 63, Appl

ALIGNMENTS

RESULT 1  
US-09-649-996-14  
; Sequence 14, Application US/09649996  
GENERAL INFORMATION:  
APPLICANT: Jono, Keith E.  
Plowman, Gregory

TITLE OF INVENTION: KINASE GENES AND USES  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" diskette, 1.44 Mb

COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSO for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/649,996  
FILING DATE: 29-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/976,255  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 229/182  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION ON SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS  
LENGTH: 1503 amino acids  
TYPE: am1503 acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-649-996-14

Query Match 8.2%; Score 76.5; DB 5; Length 1503;  
Best Local Similarity 24.7%; Pred. No. 7.3;

Matches 48; Conservative 19; Mismatches 64; Indels 63; Gaps 10;

QY 10 VTNLSQSPARGHDTLTKAGSDADGDSRSDPL-----DPRFGKDSMFYARR 59  
Db 509 VFESSLSDPGKGQD-----DSQDVPLRPGVVPVVDAAHNLVSGSDYIQLE 556  
QY 60 EASATILGLDKTSLHTDD-----ELDALFDDVQKARAVRGLVEDNESD-S 106  
Db 557 EKSGSNLELDYPPALLTMDNPRTGPELSQAL-----RSVE--LRESSTDEDF 607  
QY 107 QSSFPVRDKVMYNSGNTIDNFSLDATM-CTPRSDDDGDLAFGDKKSKDASN----VD 161  
Db 608 QSSDTPKDSLSLP-----DLHVTSGPESPNN-----IFNDVKSEDLPSHQKIFD 653  
QY 162 VEELRQQAQWERA 175  
Db 654 LMELVGQADFKA 667

RESULT 2  
US-09-792-024-76  
; Sequence 76, Application US/09792024  
; GENERAL INFORMATION:  
; APPLICANT: Roemer, Terry  
; APPLICANT: Jiang, Bo  
; APPLICANT: Boone, Charles  
; APPLICANT: Bussey, Howard  
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug  
; TITLE OF INVENTION: Targets Discovery  
; FILE REFERENCE: 10182-004-999  
; CURRENT APPLICATION NUMBER: US/09/792,024  
; NUMBER OF SEQ ID NOS: 490  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 76  
; LENGTH: 759  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-792-024-76

Query Match 7.8%; Score 73.5; DB 5; Length 759;  
Best Local Similarity 24.5%; Pred. No. 5.8;  
Matches 47; Conservative 26; Mismatches 66; Indels 53; Gaps 9;

QY 1 QEOSLLRHSVNSQLSGSPARGHDTLTKAGSDADGDSRSDPLPHFRGKDSMFYARR 60  
Db 478 EKSGLPQTGVNNVTSIPG-----IELLAKWKEQGLDSREDEE-----DQANVEYDDDE 527  
QY 61 AASAIIG-----LDQKLSHLTDD-ELDALFDDVQKARAVRGLVEDNESDSOSSFP 111  
Db 528 DADSDGWDIVDSKEINISOSDDNEDEQEPKPKGA-KTGAEDNDE-----579  
QY 112 VRKYTASGNTIDNFSLDATMPTGTPRSDGDLAFPGDKKSKDASNVYVEE--LRQQ 169  
Db 580 VSDLESS-----DQDDDEDS-----EENKQKAVADSEEPPTKKOK 615  
QY 170 AQNEALAKTAKO 181  
Db 616 IRNEADNINEQ 627

RESULT 3  
US-60-248-505-1134  
; Sequence 1134, Application US/60248505  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: c1000918  
; CURRENT APPLICATION NUMBER: US/60/248,505  
; CURRENT FILING DATE: 2000-11-15

; NUMBER OF SEQ ID NOS: 1998  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQUENCE 1134  
; LENGTH: 393  
; TYPE: PRT  
; ORGANISM: Human  
US-60-248-505-1134

Query Match 7.8%; Score 73; DB 6; Length 393;  
Best Local Similarity 22.8%; Pred. No. 2.8;  
Matches 44; Conservative 26; Mismatches 57; Indels 66; Gaps 10;

QY 3 QSLILRHSVNSQLSGSPARG-----HDTLSKAGSDADGDSRSDPL 44  
Db 48 QALLKDSIV-QCTARERPHAFLEKFEKAKQIQKAGTITSRDEISPP 105  
QY 45 P-----HFRGKDSMFYARRAAS--ALGLDOK-----TSLTDDDEL 81  
Db 106 PPNVYKRRRGCAISAEVTEADAASYVRKYDKYTKMAAKAIEKNVLSHLDNERS 165  
QY 82 ALFDDVQKARAV-RGLVEDNESDSOSSPVRQK-----YMSN-----GTIDNES 126  
Db 166 DIFDAMFSVFSFINGETVIQGDGEGD--NPFYLDQGETDVYVNNEMATSGEGSGFELA 222  
QY 127 LDTATIMEGTPRS 139  
Db 223 L-----IVGTPTA 230

RESULT 4  
US-60-248-505-737  
; Sequence 737, Application US/60248505  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: c1000918  
; CURRENT APPLICATION NUMBER: US/60/248,505  
; CURRENT FILING DATE: 2000-11-15  
; NUMBER OF SEQ ID NOS: 1998  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 737  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Human  
US-60-248-505-737

Query Match 7.7%; Score 72.5; DB 6; Length 425;  
Best Local Similarity 23.5%; Pred. No. 3.4;  
Matches 39; Conservative 24; Mismatches 52; Indels 51; Gaps 9;

QY 33 DADGDSRSDPLPHFRGKDSMFYARRAASAILGLDKLSHLTDDDELALFDD-----86  
Db 48 DIDDERHNP---HKIKSCFLPKSMGKS-----KSSH-----DLLKDDPHLSV 90  
QY 87 --VQKARAVRGLVEDNESDSOSSFPVRQY-----MSGNTIDNFSLDATMPTGT 136  
Db 91 PVVESEKGDAPLDVDDGEDESAE-----HDEYDGDKNMLHRIAKKLKDYTSANV--- 142  
QY 137 PRSDDDGDLAFPGDKKSKDASNVYVEELRQQAQWERAALAKTAKO 182  
Db 143 -KSAGEGEV-----EKKSYSRS-----EELKKEAQLARELLAAKOK 178

RESULT 5  
US-09-739-449-11627  
; Sequence 11627, Application US/09739449  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.

QY 138 SNVDVEELR 166  
| |  
| |

[illegible][illegible]

Qy	128	DTATMPGTPRSDDDGDALFFGDKKSKQDASNVDVELRQQQAQMEELAKTAKQEF	183
		:: :     :   :    :   :   :   :	
Db	475	DDSVNIPGTTSRFQG---IPEPPQSQDLNN-----SOGKOEDSTNRKKKF	520

## RESULT

```

US-09-792-024-117
> SEQUENCE 117, Application US/09792024
> GENERAL INFORMATION:
> APPLICANT: Roemer, Terry
> APPLICANT: Jiang, Bo
> APPLICANT: Boone, Charles
> APPLICANT: Bussey, Howard
> TITLE OF INVENTION: Gene Disruption Methodologies for Drug
> TARGET: 117
> FILE REFERENCE: 10182-004-999
> CURRENT APPLICATION NUMBER: US/09/792,024
> CURRENT FILING DATE: 2001-02-20
> NUMBER OF SEQ ID NOS: 490
> SOFTWARE: FastSeq for Windows Version 4.0
> SEQ ID NO 117
> LENGTH: 714
> TYPE: PRT
> ORGANISM: Candida albicans
US-09-792-024-117

```

Query Match	7.2%	Score 68;	DB 5;	Length 714;
Best Local Similarity	20.8%	Pred. No. 17;		
Matches 42;	Conservative 27;	Mismatches 67;	Indels 66;	Gaps 9;

[illegible]

## RESULT 10

```

US-09-792-024-106
? Sequence 106, Application US/09792024
? GENERAL INFORMATION:
? APPLICANT: Roemer, Terry
? APPLICANT: Jiang, Bo
? APPLICANT: Boone, Charles
? APPLICANT: Bussey, Howard
? TITLE OF INVENTION: Gene Disruption Methodologies for Drug
? TARGET:
? FILE REFERENCE: 10182-004-999
? CURRENT APPLICATION NUMBER: US/09/792,024
? CURRENT FILING DATE: 2001-02-20
? NUMBER OF SEQ ID NOS: 490
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 106
? LENGTH: 1042
? TYPE: PRT
? ORGANISM: Candida albicans
US-09-792-024-106

```

Query Match	7.1%	Score 67;	DB 5;	Length 1042;
Best Local Similarity	21.7%	Pred. No. 35;		

	Matches	31;	Conservative	33;	Mismatches	45;	Indels	34;	Gaps	6;
Qy	56	YAREAA	SAILGDKQIKSHLTDDLDALFDVQKA---	--RAVRGLVEONEDSDQSFPV	112					
		: : :	: : :   :   :	: : :   :   :	:	:	:	:	:	:
Dd	914	YFRVLKSAI---	EESMPDENNDAMNGKETEQIDYKAIEQLQDLEAGDAEIAEM	970						
		: : :	: : :   :   :	: : :   :   :	:	:	:	:	:	:
Qy	113	R-----	-DKY--MNSGTINDFSLDTATMTGTPRSDDGDFALFFDKKKSQDAS	158						
		: : :	: : :   :   :	: : :   :   :	:	:	:	:	:	:
Dd	971	REKORELINALNDKYAETAEADWEKSMDKATKGKNVVS-----	-IKSKRKSENAN	1024						
		: : :	: : :   :   :	: : :   :   :	:	:	:	:	:	:
Qy	159	NVDVELROOQAOMEALKTAQ	181							
		: : :	: : :   :   :	: : :   :   :	:	:	:	:	:	:
Dd	1025	DIVEKEMK-	AVKSKK	1039						
		: : :	: : :   :   :	: : :   :   :	:	:	:	:	:	:

RESULT 11

```

RESULTS:
; Sequence 8477, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkie, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 8477
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
; US-09-739-449-8477

```

Query Match 7.18; Score 66.5; DB 5; Length 529;

	Best Local Similarity	25, 28; Pred. No. 16;	Matches	32; Conservative	19; Mismatches	45; Indels	31; Gaps
QY 44	LPHFGRKDSWFYAR--REASAILGLDOOKTSHLTDELDAIPDDVQKARVARRGVLVEDN	101					
Db 221	LPYL-----DGFVAREVDAASATTSIVS-----TGADYTDVSYSDLERLQNPKLAVE--	270					
QY 102	EDSDQSSEFPVRKDYMSNGTIDNFSIDATIMFGTPR-----SDDGDGALPFQDKKS	153					
Db 271	-----VVTDAVLNNAALFEENLENAAILAKKEVRLAHALADINDFINDAIFFG--TA	319					
QY 154	KQASNV 160						
Db 320	KPAGSNI 326						

**RESULT** 12

```

RES001 14.
US-09-739-449-11845
; Sequence 11845, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hixie, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11845
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11845

```

Query Match 7.1%; Score 66.5; DB 5; Length 606;

```
Query Match
Best Local Similarity 24.1%; Score 66; DB 5; Length 812;
Matches 34; Conservative 17; Mismatches 53; Indels 37; Gaps 6;

QY 62 ASATLGQKISHLTDELDFDVOKARVRGLVEDNEDSSOSSFPVRKYNSTG 121
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 426 APNLLIDEPHNLIDIS-----RRALTEALNIOYSGAVILISHORLLEAT 471
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 TDNPSL---DTAFTWCT-----PRSDGDALFFGDKSKQDASVYVELR 166
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 472 VORLWLVADCTVKTFEGDMEERYDLVSSGRKKDKDTTEA--AQTSKADQRANAE-R 528
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 OQAOQMEELK-----TAKOE 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 529 AQLAPLKKKINIESLTAKLE 549
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-739-449-12024
; Sequence 12024, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-1015490JC
; CURRENT APPLICATION NUMBER: US/09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US/09/514,000
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 12024
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12024

Query Match
Best Local Similarity 7.0%; Score 66; DB 5; Length 684;
Matches 36; Conservative 24; Mismatches 64; Indels 38; Gaps 7;

QY 26 TLSKAGSDAG--DSRSDSPLPIHFGRKSDNFWYARREASAILGLQKISHLTDELDAI 83
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 TKYKENEAEENRGATGCPILLD-----SD-----DAVKMKTAKKRGVYTWELNS- 51
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 84 FDDVOKARVRGLVEDN-----RSDSSOSSFPVRKYNSTGTDNFSLOTAL 131
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 ---VLPSEVTSEQIEDYMAHLSMDGNGINVEDAEAEAFED---DGDSDNEESE--- 101
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 132 TWPCTPRSDDDGDALFFGDKSKQDASVYVELRQOQAOME 173
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 102 ---GGLAPSGGTALATAKKKEPTDRTDDPVRMYLREMGVSVE 140
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-739-449-9492
; Sequence 9492, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-1015490JC
; CURRENT APPLICATION NUMBER: US/09/739,449
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US/09/514,000
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9492
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9492
```

```
Query Match
Best Local Similarity 18.8%; Score 66; DB 5; Length 812;
Matches 37; Conservative 36; Mismatches 68; Indels 56; Gaps 7;

QY 16 GSPAPGHRDRTLKAGSDA-----DGSRSD-----SPLPHFRKQSD 53
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 426 GTLAAGTPTTPTTTLGATTTGSGVTHYNDADAKAKALSLFKSISKVTYVMDPLGTNAV 485
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 54 WYKARASASAILGLQKISHLTDELDFDVOKARVRGLVEDNEDSSOSSFPVR 113
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 486 YTYALR-----AVGLDPKAVNISD-----IQRVLSDLQDKSYVYTLKD 525
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 114 DKYNSGTIDNFSLOTATWPCTPRSDDGDAL-----FFGDKSKQDASVYVEE 164
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 526 DRYVLAELFNFTKDGSGSPILAQSETELTQMSADYIKKSAFGTEKDEAAQT-----E 581
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 165 LRQOQAOQMEELKATAKQ 181
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 582 AKYFSGEMQK-IKTLSE 597
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-09-822-246-2
; Sequence 2, Application US/09822246
; GENERAL INFORMATION:
; APPLICANT: HENKUDJOY, Kennedy et al.
; TITLE OF INVENTION: IDENTIFIED HUMAN TRANSPORTER PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001149
; CURRENT APPLICATION NUMBER: US/09/822,246
; CURRENT FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Human
US-09-822-246-2

Query Match
Best Local Similarity 7.0%; Score 65.5; DB 5; Length 727;
Matches 25; Conservative 15; Mismatches 46; Indels 15; Gaps 4;

QY 73 SHLTDELDFDVOKARVRGLVEDNEDSSOSSFPVRKYNSTGTD----NFSL 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 SRFDQEDD---DDYTPAGTNYNGEANDDESGSEATEGHDDEDIYEGEYOGIPSMNOAK 103
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 128 DTAITWPGTPRSDDGDALFFGDKSKQDASVYVDELRRQQ 168
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 104 DSTVSV-GQPKGDE-----YKDRRELESERRADEELAEQAQ 137
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: April 25, 2001, 10:16:34
Job time: 305 sec
```





Result	Query No.	Score	Match	Length	DB	ID	Description
1	4030	100.0	784	20	Y06618		Thermoplasma
2	1668.5	41.4	1816	21	B36227		Human kine
3	1558	41.1	1103	21	Y51328		Human KLIM1
4	1778.5	31.7	504	21	B63189		Gene 5 human
5	1276	31.7	503	21	B63190		Human hsc
6	814	20.2	955	15	R57365		k39 polyprotein
7	814	20.2	955	17	W03361		Leishmania
8	662.5	16.4	2954	20	Y01632		Amino acid
9	649.5	16.1	975	19	W72746		Drosophila
10	625	15.5	411	19	W72745		Drosophila
11	625	15.5	411	19	W72744		Drosophila

## ALIGNMENTS

diagnosis of fungal infection and neurodegenerative disease

XX Claim 5: Page 70-71: 75pp; English.  
 XX This sequence represents Thermomyces lanuginosus TL-gamma, a novel  
 XX ATP-dependent, plus end-directed microtubule motor. Protein that is  
 XX a member of the tau-104 family and contains a superfamily. The  
 XX invention provides TL-gamma nucleotide acid seq. X87656, proteins  
 XX and antibodies, and methods of screening for TL-gamma mutants  
 XX potentially useful for treating hyphal fungal infections and  
 XX diseases caused by mutated TL-gamma, e.g. neurodegeneration  
 XX involving anterograde axonal transport, such as Alzheimer's  
 XX Parkinson's or Huntington's diseases or amyotrophic lateral  
 XX sclerosis. Detection of TL-gamma allows differentiation between  
 XX hyphal and non-hyphal fungal infections.  
 XX Sequence 784 AA;  
 XX  
 Query Match 100.0%; Score 4030; DB 20; Length 784;  
 Best Local Similarity 100.0%; Pred. No. 4e-300;  
 Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MSGGGNIKVVVRFPNAREIDRGAKCIVRMEGNQITLTPPGAEEKARKSGKTMDGPK 60  
 DB 1 msgggnikvvvrpfnareidrgakeivrmegnqitltppgaeeakrsgktindgpk 60  
 OY 61 AFADRSYWSFDKNAPFAQEDLPQDLGVPLLDNAFKGNKNCIFAYGTGSGKSYSHAG 120  
 DB 61 afadrsywsfdknayvarqedlqdlgvplldnafkgnncifaygtgsgksysmag 120  
 OY 121 XGEHGVIPRICODMFRINELQOKNUTCTVEVSYLEIYNERNVRLNPNSTGKLNKVR 180  
 DB 121 ygehgvipricdmfrinelqkdnktctvevsyleiynervrldlnpstgknkvr 180  
 OY 181 HPSTGPGYVEDLAKLVVRSFOEINLMDGNKARTVAATNNNETSRSHAVFTLTQLKWH 240  
 DB 181 hpstgpgyvedlaklvvrsfgeienlmdgnkarktvaatnnnetssrshavftltqlkwh 240  
 OY 241 DEETKMDTEKAKISLVLAGSERATSGTAGRLKEGAEINRSLSLTGLRVIAALADMS 300  
 DB 241 deetkmdtekakislvlagseratstgtarlkageaieinrslstglrviaaladms 300  
 OY 301 GKOKKQLVPYRDSVLTLWKLSDSGNSMTAMIAISPADINFEETLSTLRVADSAKR 360  
 DB 301 gkqkknqlvpyrdsvltlwlkdsdgnsmtamiaispadinfeetlstlrvyadsakrik 360  
 OY 361 NHAVVNEDPNARMIRELKEEALORSLQSGGGGGAGGSGGVPVEESYPDTPLEKQIV 420  
 DB 361 nhavvnedpnarmirelkeelaqirsklqsgggggagsggvpveesypdtplekqiv 420  
 OY 421 SIQPDATVTKMSKAEIVQELNOSSEKLYRDLQNTWBEKLAETEEHKEEAAELGISI 480  
 DB 421 siqpdattvkmskaeivqelngseklrydlngtweeklakceelhkeeraaelgisi 480  
 OY 481 EKGVGPGYHSEKMPHLVNLSDPLAECLVYNIKPGQTRGVGNVNDTQAEIRLNGSLIK 540  
 DB 481 ekgvgpgyhsekmphlvnlstdplaeclvynikpgqtrgvgnvndtqaeirfngsklik 540  
 OY 541 EHCFTFENVDNVTVPNEKAVNVGVNRIDKPTRLRSGYRIITLGDHFIRFNHPPEARAE 600  
 DB 541 ehcftfenvdntvpnekaavnvgrvridkptrlrsgyriitlgdhfrfnhppearae 600  
 OY 601 RQEOSLLRHVSNSQLGSPAGRHDTLSKAGSDAGDSRSPLPHFHKGDSDFYARR 660  
 DB 601 rqeosllrhvsnsqlgspagrhdtlskagsdagdsrdsplphfrkgsdwyfarr 660  
 OY 661 FAASATLGLDQKLSHTDELDALFDVQVAKARVGLVEDNEDSDSGSPFVRDYKSN 720  
 DB 661 faasatlglqklshtdelalfdvqvakarvrglvednedsgsfvrdykn 720  
 OY 721 GTDNPSLDPTATFMTGCTPRSDDDGALFDGCKSQDASNDVYELPQQOQMEALKA 780  
 DB 721 gtdnpsldptatfmgctprsdddgalfdgcksqdasndvyelpqqoqmealka 780

OY 781 QKEF 784

|||||

DB 781 Kqef 784

|||||

RESULT 2

B36227

ID B36227 standard; Protein; 1816 AA.

XX B36227;

XX B36227;

XX B36227;

DT 19-FEB-2001 (first entry)

XX

XX Human kinesin-like protein HKLP SEQ ID NO: 4.

XX

XX Human: kinesin-like protein: HKLP; KIF1; cell division; cancer;

XX intracellular transport; neurological disorder; infertility;

XX biallelic marker; spontaneous abortion; neonatal chromosome disorder;

XX aneuploidy.

XX Homo sapiens.

XX WO200063375-A1.

XX 26-OCT-2000.

XX 20-APR-2000; 2000WO-IB00562.

XX 20-APR-1999; 99US-0130217.

XX (GEST ) GENSET.

XX Bougueleret L, Dufaure-Gare I, Grel P;

XX WPI; 2000-665242/64.

XX N-PSDB; C66550.

XX An isolated or purified human kinesin-like protein (HKLP) encoding

XX polynucleotide used to detect HKLP polynucleotides in a sample

XX comprises a contiguous span of at least 12 nucleotides .

XX Claim 46: Page 189-192; 199pp; English.

XX

XX The present invention describes the coding and protein sequences of the

XX human kinesin-like protein HKLP. It is thought that the protein could be

XX involved in neurological disorders, infertility, spontaneous abortion,

XX neonatal chromosome disorders, aneuploidy and cancers. This is due to its

XX function in the movement of microtubules. The protein shows homology to

XX the murine KIF1A and KIF1B proteins. The sequences disclosed in the

XX invention can be used in the isolation of similar human proteins and in

XX vector production. In addition, the biallelic markers shown can be used

XX in disease diagnosis and population studies.

XX Sequence 1816 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 181 HPSTGYVEDLAKLVRSFOEINLMDGNKARTVAATNMNSTRSSSHAVFTLTITLQKWH 240  
 Db 171 hpllgpyvedlslavtysdiadlmdgknkartvaatnmnstrssshavftlvtqkhh 230  
 QY 241 DEETKMDTEFKVAKISLVLDLAGSERATSGATGARKKEGAETNRSLSSTGRVIAALADNS-- 298  
 Db 231 dnetnlssekvskisvdlagseradstgkgrtkeganinkeltlgtkvksalaeavdn 290  
 QY 299 ---SSGQKKKQOLVPRDSYLVTLKDSLGNSWTAMTAALSPADINFEETLSLRVADS 355  
 Db 291 ctskxkxktcfpydsvytlwlenlgnsrtamlaalspadlnfeetlsrlyadr 350  
 QY 356 AKRIKNHVVNEDPNARMTELKELAQRLSKLOSSGGG-----GGAG----- 399  
 Db 351 akqlcnaevnedpnarivelkeevtrikdlraaglgldldplddyegsgskylk 410  
 QY 400 -----GSGGPVEESYPPDTPLEKO-----IVSIQQPDATVKK 431  
 Db 411 dfqnkhryllasenqrphfstaemsgsltas-pascslssqvgltsvtslq-erlmet 467  
 QY 432 MSKAEIVDOIJOSEKLYRDLNMTWEKELAKTEETIKHEREAALELGISTEK--GFVGYPH 489  
 Db 468 pggaealerikesekielnetweekirktealmerceallaemgvalredggltgfs 527  
 QY 490 SKEMPHLVNSDDPLAELVYHVKPQTVGVNNDTQAEITRNGSKILKEHCTFNW- 548  
 Db 528 paktphlvlnedpimsecllylkgdltvrgqadeerrqdlvisgahkeencilfrser 587  
 QY 549 ----DNVYTVPEKAAVNGVRLDKPRLRSYRITLGDHFIFENPEEABARQEO 604  
 Db 588 ansgevitclepcarsetyngkrvcpqlrgsnrlimgknhvfrfhhpegrarek- 646  
 QY 605 SILLRHSVNSQLGSPAGCRHRTLSKAGSDADGSRSDPLPHERGKGSDFWYARREAS 664  
 Db 647 -----tpsaetpaepvdtfaqrelle 668  
 QY 665 AILGLDOR-----ISHLTDDELDAIFDD-----VQKRAHVRGLVED 701  
 Db 669 k-agldmkemektqlqemelllytkeeeadilleqridyeskqlqkqvtrslaaet 727  
 QY 702 NEDSDQSQSPF 712  
 Db 728 teeeeeeavp 738

RESULT 3  
 ID Y51328 standard; Protein; 1103 AA.  
 AC Y51328;  
 DT 17-APR-2000 (first entry)  
 DE Human KLIMP protein.  
 KW KLIMP; kinesin-like motor protein; cytoskeletal; anticonvulsant; human;  
 KW anti-Alzheimer; anti-Parkinsonian; antidiabetic; anti-ulcerative; cancer;  
 KW immunomodulatory; antiinflammatory; anti-AIDS; antirheumatic; treatment;  
 KW antiarthritic; diagnosis; neurological disorder; vesicular transport.  
 OS Homo sapiens.  
 UN US6013454-A.  
 PD 11-JAN-2000.  
 PF 28-SEP-1998; 98US-0162373.  
 PR 28-SEP-1998; 98US-0162373.  
 PA (INCY-) INCYTE PHARM INC.  
 XX Tang YT, Corley NC, Patterson C, Guegler KJ;

XX WPI: 2000-126064/11.  
 DR N-PSOB: 244744.  
 XX Nucleic acid sequences encoding a human kinesin-like motor protein (KLIMP) useful for the treatment of diseases associated with inappropriate KLIMP expression such as cancers, neurological disorders and disorders of vesicular transport -  
 Claim 1: Fig 1A-J; 38pp; English.  
 XX This invention describes a novel human kinesin-like motor protein (KLIMP) which has a novel amino acid sequence, and which has anti-inflammatory, anti-AIDS, antirheumatic and antiarthritic activity. (1) and the protein it encodes may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate KLIMP expression such as cancers, neurological disorders and disorders of vesicular transport. For example, (1) (and vectors containing (1) (iv)) and the KLIMP polypeptide may be used to treat disorders associated with decreased KLIMP expression such as cancers (e.g. lymphoma, melanoma and cancers of the breast lung and prostate), neurological disorders (e.g. epilepsy, Alzheimer's disease and Parkinson's disease), disorders of vesicular transport (e.g. diabetes mellitus/insipidus, Grave's disease and gastric/duodenal ulcers), and some immune/inflammatory diseases (e.g. acquired immune deficiency syndrome AIDS), rheumatoid arthritis and toxic shock syndrome). This sequence represents the human KLIMP protein described in the method of the invention.  
 XX Sequence 1103 AA:  
 QY 411: Score 1658; DB 21: Length 1103;  
 Query Match 47.3%; Pred. No. 3.7e-118;  
 Matches 353; Conservative 120; Mismatches 159; Indels 114; Gaps 16;  
 QY 4 GGNKVVVRVPPNARIEDRGAACIVRMENQOTILTPPGAESKARKSGKTIMDGPKAFA 63  
 Db 3 gasvkavvrpfnaretsqadkcvvsnqnttislmp-----kgsdkapksft 51  
 QY 64 FDRSYMSFKNA-PNYARQEDFDOLGVPFLDNLAFKGNCFAYGOTGSGKSYSMWGYG 122  
 Db 52 fdysywshtstedqfasqggyvrdigeemllhafegynvcifaygtqgagkytmnqrq 111  
 QY 123 K--EHGVIPRICQDMFRINELQNDKNTCTVEVSYLEIYNERNVRLDNLNPNSTKGNLKVRE 180  
 Db 112 epvgqglvpqldedfivrsfoeintlmdgkarkartvaatnmnstrssshavftlvtqkhh 170  
 QY 181 HPSTGYVEDLAKLVRSFOEINLMDGKARTVAATNMNSTRSSSHAVFTLTITLQKWH 240  
 Db 171 hpllgpyvedlslavtysdiadlmdgknkartvaatnmnstrssshavftlvtqch 230  
 QY 241 DEETKMDTEFKVAKISLVLDLAGSERATSGATGARKKEGAETNRSLSSTGRVIAALADNS 300  
 Db 231 dnetnlssekvskisvdlagseradstgkgrtkeganinkeltlgtkvksalaeavdn 290  
 QY 301 GQKKKQOLVPRDSYLVTLKDSLGNSWTAMTAALSPADINFEETLSLRVADSARKIK 360  
 Db 291 -hkrksdfpydrsvlwlknlgnsrtamlaalspadlnfeetlsrlyadrckqir 349  
 QY 361 NHAVVNDPNARMTELKELAQRLSKLOSSGGGAGG----- 400  
 Db 350 cnaifnedpnarivelkeevtrikdlraaglgldldplddyegsgskylk 410  
 QY 401 -----SGGPVEESYPPDTPLEKO-----IVSIQQPDATVKKKAEIVQKSEKLYRDL 451  
 Db 410 pvspsptthgealqpsfnt--esql-----gpeameriqetektlael 454  
 QY 452 NOTWEKELAKTEETIKHEREAALELGISTEK--GFVGYPHSEKMPHLVNSDDPLAELV 509  
 Db 455 netweekirktealmerceallaemgvalredggltgvisgahkeencilfrser 587  
 QY 510 VYNIKPGQTRVGNVDQAEITRNGSKILKEHCTFNW-----DNVYTVPEKAAVW 564

Db 515 lylhkgvtrvgvdm-----lklgtgfrehclfrispqdgdevvvtlcepgatyv 570  
 Qv 565 NGVRIDKPTRLSRYRIILGDHFHFRFNHPFARERQSLRHSHVTSQSGSPAGRH 624  
 Db 571 ngklvteplvlsnrvimgknhfrfnbeqarler-----gvppp--- 614  
 Qv 625 DRTLXSGSDADGSRSDPLPHFRGDSOWFYARRRASAAILGLDQIKSHULTDELDA 684  
 Db 615 -----pyppepvdwnfaqqelleq-ggidikie--mekrlgl 650  
 Qv 685 FDDVQKARAVRGVLVDNE---DSDS 707  
 Db 651 endyrkekeadlillleqgriyadsds 676  
 RESULT 4  
 ID B63189 standard; Protein; 504 AA.  
 AC B63189;  
 DT 26-MAR-2001 (first entry)  
 XX Gene 5 human secreted protein homologous amino acid sequence #115.  
 DE Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasia;  
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
 KW angiodenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; wound healing; skin aging;  
 XX Drosophila melanogaster.  
 OS WO2000061629-A1.  
 PN 19-OCT-2000.  
 PD 06-APR-2000; 2000WO-0909071.  
 PE 09-APR-1999; 99US-0128694.  
 PR 20-JAN-2000; 2000US-0176931.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 Qv Ruben SM, Komatsoulis G;  
 XX WPI: 2000-647420/62.  
 Isolated nucleic acid molecule encoding a human secreted protein is  
 used in preventing, treating or ameliorating a medical condition -  
 XX Disclosure; Page 487-488; 533pp; English.

XX F23273 to F232421 encode the human secreted proteins given in B63134 to  
 CC B63182, B63183 to B63231 represent more human secreted proteins and  
 CC polypeptides homologous to them. Human secreted proteins have activities  
 CC based in the tissues and cells the genes are expressed in. Examples of  
 CC activities include: immunosuppressive, antiarthritic, antirheumatic;  
 CC antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
 CC neurotropic; neuroprotective; antibacterial; virucide; fungicide; and  
 CC ophthalmological. Polyclonal antibodies and proteins can be used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also  
 CC used in diagnosing pathological conditions or susceptibility to a  
 CC pathological condition. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders

CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiodenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain tissues before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities. F22364 to F23372 and B63133  
 CC represent sequences used in the exemplification of the present invention.  
 XX Sequence 504 AA:

Query Match 31.7%; Score 1278.5; DB 21; Length 504;  
 Best Local Similarity 50.9%; Pred. No. 1.3e-89;  
 Matches 274; Conservative 77; Mismatches 144; Indels 43; Gaps 6;

Qv 59 PKAFADFDSYMSFDKNAPYARQEDLFQDLGVPLDINAFKYNMCIFAYGTGSGKSYSM 118  
 Db 8 pktfaifdhfyslnpidentfasqetdcvgrgldnafggnacifayggtgsgksytm 67  
 Qv 119 MGYKKEHGVPRIICODMFRINELQKNICTCTVEVSYLEIYNERVRDLNPS-TKGNLK 177  
 Db 68 mgtqeskgilpricdqlfsalan-kstpelmykvevsymelynekvhdldpknqsklk 126  
 Qv 178 VREHPSTGPVYEDLAKLVRSFOEINIMDEGNKARTVAATNNNETSSSRSHAVFTLTQ 237  
 Db 127 vrehnmvpyvgdglavtsygdldnImtegnksrtvaatnmaesrshavsvltq 186  
 Qv 238 KWHDEETKMDTEKAKISLYDLASERATSTGATGABKEGAEINRSLTGLRVIAALAD 297  
 Db 187 lltcdgtvgeksrmlvdlageravktgdvgrlkegnslntltglvsklad 246  
 Qv 298 MSSGKKKN-QLVPRDSVLTLLKSLGNSMTAMIAISPADINFEETSTLTVADSA 356  
 Db 247 qsnkgksgndkfpyrdsvitllkldngnartvmvatispsadnyetstlrvyadra 306  
 Qv 357 KRKHAVVNEPDNARMIRBELKEALQRLSKLQSSGGGGGAGSGGVPVEESPYPOTPLE 416  
 Db 307 krivnhavvnedpnariirelthevetrlsmikha-----tgspsy----- 346  
 Qv 417 KQIVSIQPDATYKMSKAEIVQUNSEKLYRDLNQTWEKLAETETHEKREAALEEL 476  
 Db 347 -----gdvgkkaesenlmkqisqtweeklvkterlqnerqgalekm 388  
 Qv 477 GISTEKGFVGYHSEKEMPHLVNLSDDPLAELVYNNKPGQTRGVNVDQAOEIRLNS 536  
 Db 389 glsvqas--gikveknkyvlvnladpslnellvyikdrtlilgrtligsggpdglsl 446  
 Qv 537 KILKECHTFENVNDNVTVIPNEKAAVWVGVDRDKPTRLRSGYRIILIGDFHFRNHP 594  
 Db 447 glqpehcvtltedsglymepvggarcfcvngsaavektplqngdrillwgnhhfrvns 504

RESULT 5  
 ID B63190 standard; Protein; 503 AA.  
 AC B63190;  
 DT 26-MAR-2001 (first entry)

XX Human secreted protein sequence encoded by gene 5 SEQ ID NO:116.  
 DE Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasia;  
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
 KW angiodenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; wound healing; skin aging;

[illegible]

DR	NFI; 1994-249402/30.
DR	N-P5DB; Q70152.
PT	
XX	Diagnosis of Leishmaniasis - by determining the presence of
PT	antibodies that bind to a K39 repeat unit antigen
XX	
PS	Disclosure; Page 12-15; 28pp; English.
XX	
CC	The K39 polypeptide comprises a number of repeated units (described
CC	in R57366). Detection of antibodies directed against this repeated
CC	unit in a patient's sample is indicative of leishmaniasis. The
CC	antigenic repeat unit can itself be used as a vaccine to protect
CC	against infection by a leishmania parasite.
XX	
SQ	Sequence 955 AA;
	Query Match 20.2%; Score 814; DB 15; Length 955;
	Best Local Similarity 28.1%; Pred. NO. 1-4e-53;
	Matches 23; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

```

OY 7 IKVVVRVFPFNARE--IDRGAKCIVRMEGNQTILTPP-----PGAEKARKSGKTIMD 57
Db 13 VKSVVRPLNERENNAPEGKTVAAKQAQAAVVTVKVIGGSSNGSAESMGTAIRVAQD 72
OY 58 GKPAFAFDRSTWSF--DKNAPYARQEDLFQDLGVPLLDNAPKGYNWCIFYAGTGSCK 114
Db 73 ----fQdhvfwsvetdaagcatpatqadvrtigylvqhafgnscifayqgtsqk 128
OY 115 SYSWMG-----YGKEHGVIPRICODMFRINELQKDKNLTCTVEYSLEYINVRDILL 168
Db 129 tytmgadvlsalgegnvpriciefarkasveaghsrwlvelgyevynrvsdll 188
OY 169 NPSTKG-----NLKVREHSTGPVVEDLAKLVRSFOETENLMDGNKARTVAATNM 220
Db 189 gkrkkgvkggeevyvdrehpsrgvflgeqrlvegslddvrlieignvritastkm 248
OY 221 NETSSSHAVFTLTL----TQKWHDEETKMDTEKVKAKISLVDLAGSERATSTGATGARLK 276
Db 249 ndrstrshaiimllireertmttksetirtagkssrmlnldvlgsvagveggqtk 308
OY 277 EGAEINRSLTGLRVTAALADMSS--GKQKNQLVPRDSVLTWLLKDSLGGSNTAMIAA 335
Db 309 eathnlsitlgrvidmadmatgkagysvapfrdsklflfkdsiggnsktmiat 368
OY 336 ISPADINFEETLSTLRYSADSAKRINHAHVNEDPNARKIRKELKEELAQRLSKLOSSGGG 395
Db 369 vspsainyeetlstrysardinvvaqnedprarrirleeeqmedmrqam----- 421
OY 396 CGAGSGGSPVEE-----SYPPDPPLK 417
Db 422 --agdpayvseikxlallesaqkraadlqalererhngvqerllrateaekseles 479

```

## RESULT 7

ID W03691 standard; Protein; 955 AA.

XX W03691;

XX AC

XX DT 09-MAR-1997 (first entry)

XX Leishmania chagasi

XX Leishmania chagasi; acidic ribosomal antigen; LcP0;

XX epitope; K39.

```

XX Leishmania chagasi.
XX W09633414-A2.
XX 24-OCT-1996.
XX 19-APR-1996; 96WO-US05472.
XX 21-APR-1995; 95US-0428414.
XX (CORI-) CORIAX CORP.
XX Reed SG;
XX WPI; 1996-485884/48.
XX N-PSDB; T42166.
XX New Leishmania acidic ribosomal p-protein family poly:peptide - used
XX to develop prods. for diagnosis, detection and protection against
XX Leishmania infections
XX Disclosure; Page 36-43; 76pp; English.
XX Compounds including polypeptides that contain at least an epitope of
XX the L. chagasi acidic ribosomal antigen LcP0 are useful in a variety
XX of immunoassays for detecting Leishmania infection. Portions of
XX LcP0 (T42164) contg. at least the 17 C-terminal amino acids (T42165)
XX have been found to generate a signal in an ELISA that is equivalent
XX to that generated by the full length LcP0. A combination
XX of polypeptide may also be used, comprising an LcP0 epitope along with
XX the K39 repeat unit antigen having the sequence given in W03690.
XX Sequence 955 AA;
XX

```

Query Match 20.2%; Score 814; DB 17; Length 955;

Best Local Similarity 28.1%; Pred. NO. 1.4e-53;

Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

```

OY 7 IKVVVRVFPFNARE--IDRGAKCIVRMEGNQTILTPP-----PGAEKARKSGKTIMD 57
Db 13 VKSVVRPLNERENNAPEGKTVAAKQAQAAVVTVKVIGGSSNGSAESMGTAIRVAQD 72
OY 58 GKPAFAFDRSTWSF--DKNAPYARQEDLFQDLGVPLLDNAPKGYNWCIFYAGTGSCK 114
Db 73 ----fQdhvfwsvetdaagcatpatqadvrtigylvqhafgnscifayqgtsqk 128
OY 115 SYSWMG-----YGKEHGVIPRICODMFRINELQKDKNLTCTVEYSLEYINVRDILL 168
Db 129 tytmgadvlsalgegnvpriciefarkasveaghsrwlvelgyevynrvsdll 188
OY 169 NPSTKG-----NLKVREHSTGPVVEDLAKLVRSFOETENLMDGNKARTVAATNM 220
Db 189 gkrkkgvkggeevyvdrehpsrgvflgeqrlvegslddvrlieignvritastkm 248
OY 221 NETSSSHAVFTLTL----TQKWHDEETKMDTEKVKAKISLVDLAGSERATSTGATGARLK 276
Db 249 ndrstrshaiimllireertmttksetirtagkssrmlnldvlgsvagveggqtk 308
OY 277 EGAEINRSLTGLRVTAALADMSS--GKQKNQLVPRDSVLTWLLKDSLGGSNTAMIAA 335
Db 309 eathnlsitlgrvidmadmatgkagysvapfrdsklflfkdsiggnsktmiat 368
OY 336 ISPADINFEETLSTLRYSADSAKRINHAHVNEDPNARKIRKELKEELAQRLSKLOSSGGG 395
Db 369 vspsainyeetlstrysardinvvaqnedprarrirleeeqmedmrqam----- 421
OY 396 CGAGSGGSPVEE-----SYPPDPPLK 417
Db 422 --agdpayvseikxlallesaqkraadlqalererhngvqerllrateaekseles 479

```

OY 418 QIVSIQ-----PQATVKKSAEIVEQL 441  
 Db 480 aaalqemtatraqkqpalnlrlseaqateralkkmaakdaalstkrirrkdaela 539  
 OY 442 NOSEKLYPOLNQTWEKLAETIEHKERAALBELGJISIEKGFVGVPSKEMPHLVNSLD 501  
 Db 540 serek1-----estvaqlreqrevaldal-----qtbqrklqaelesae 581  
 OY 502 -----DPLLAELCLVYNIKPOOTRVGNVNDQTQAEIRLNGSKILKEHCTFENVDNVTV 555  
 Db 582 rtaaedrlqlq--lclqserqtqlsylvdreltr-dlqrlgyeygetelardvalca 638  
 OY 556 PRE-----KAAVWNGVRDKPRRUSRYRIIGDPHIFRPHPEARAEQBSLRHS 610  
 Db 639 aqemearryhaavrhqltellelacedalr-----eraiaerdeaaaa-- 682  
 OY 611 VTNSQLSGPAPGCHDRTLSVAGSDADGSDRSPLDHPFKGSDWFPYARRASAILGID 670  
 Db 683 -----lidaastagq-----aresacerlslsle 705  
 OY 671 QKISHLTDBDELALFDQOKARAVRGLVEDNESDSQSFPVRDKYMSNTIDNFSLOT 730  
 Db 706 qqlre-seeraaelasqlaataaaksaaeqdrentratleqqles-----earaael 757  
 OY 731 AITMPTGTPRSDDGDAIFPGDKSKQSDASNV-----DVEELRQQAOMEELAKTKAR 781  
 Db 758 aaqlaataa-----kmaaeqdrentratleqqldraeeraaelasqlaesttak 807

RESULT 8  
 TD Y01632 standard; Protein; 2954 AA.  
 X1 X1  
 AC Y01632;  
 DT 22-JUN-1999 (first entry)  
 XX  
 DE Amino acid sequence of centromere-associated protein-E (CENP-E).  
 KW CENP-E; centromere-associated protein-E; ATPase activity;  
 KW plus end-directed microtubule motor activity; chromosome congression;  
 KW microtubule binding activity; chromosome movement; mitosis;  
 KW cell proliferation; tumor; metastasis; vascular malfunction;  
 KW inflammatory disease; immune disease; angiogenesis; hypertension;  
 KW restenosis; fungal infection; selective herbicide; fungicide;  
 KW insecticide; plant growth regulator; activator; cancer cell marker.  
 XX Xenopus sp.  
 XS W09913061-A1.  
 XX  
 PD 18-MAR-1999.  
 XX  
 PF 10-SEP-1998; 98WO-US19231.  
 XX  
 PR 11-SEP-1997; 97US-0058645.  
 XX  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Cleveland DW, Goldstein LSB, Sakowicz R, Wood KW;  
 XX NPI; 1999-229233/19.  
 DR N-PSDB; X26819.  
 XX  
 PT Centromere-associated protein-E and related nucleic acid  
 XX  
 XX Claim 5; Page 66-67; 77pp; English.  
 XX  
 CC The present sequence represents CENP-E (centromere-associated protein-E)  
 CC of Xenopus. The protein has at least one of plus end-directed  
 CC motor activity. ATPase (adenosine triphosphatase) activity and  
 CC microtubule binding activity. CENP-E is the motor that powers chromosome  
 CC movement toward microtubule plus ends and is essential for congression

CC of chromosomes during mitosis. Modulators of CENP-E can thus control  
 CC cell proliferation. Agents that modulate CENP-E activity are lead  
 CC chemopreventive, biogeriatric and diagnostic agents, e.g. for treatment  
 CC of cancer. The protein has many other functions, including angiogenesis and  
 CC metastasis; vascular malfunction; inflammation and immune diseases;  
 CC angiogenesis; hyperextension; restenosis; and fungal infections), also as  
 CC plant-protection agents (selective herbicides, fungicides and, also as  
 CC insecticides) and plant growth regulators or activators for improving  
 CC yields. CENP-E is also a diagnostic marker for dividing cells, including  
 CC cancer cells.  
 XX Sequence 2954 AA:

Query Match 16.4%; Score 652.5; DA 20; Length 2954;  
 Best Local Similarity 26.8%; Pred. No. 3,le-41;  
 Matches 226; Conservative 147; Mismatches 249; Indels 227; Gaps 29;

OY 1 MMSGGNKIVYRVPFRNAREIDRGNACIVRMEGNQITLPPPGAEKARKSKOTIMQPK 60  
 Db 1 msegdgvkvrvrplqre-----sgdqanl-----qvkagntisqvgdtk 43  
 OY 61 AFAPDSVSTFDKAPYARQEDLFDGLGVPLDNNFKNYNCIFAYGQTGSGKSYMGG 120  
 Db 44 snfdrrvfnshests-----qiygeiavplirsalqgyngtifaygqtssgkytumm 96  
 OY 121 YKRGHGVIPRICQDMFRINELQDKNLTCTVEVSYLEIYNERVDDL-NPSTKGNLVR 179  
 Db 97 tpnslgllpqaigevfk1qel--pnreflrvaymelynetvkdilcdrrfkpleir 153  
 OY 180 EHPSTGYPVEDLAKIVRSFQIEINLMDEGNKARTVAATNMNETSSRSHAVFTLITOKW 239  
 Db 154 edfnrvvyaditceelvmpvpehvlqwkgeknhygetckmhdmsrshftlmveard 213  
 OY 240 HDEFTKMD-----TEVAKISLVLDAGSRATSTGATGALFKGAEIHSLSLGRVIAAL 295  
 Db 214 rldptnscndgavmvshlnldlagaesvqtqgeyvrlegculnslsflgylvklk 273  
 OY 296 ADMSSGKQKQKQVLQVPSVFTWLLKDSLGNSMTAMIAALSPADINPFTLSTLVAQDS 355  
 Db 274 sdqag-----gfinyrdsklrlqlsngnakvtictlp--vafdetlslqfast 326  
 OY 356 AKRKINHAIVNE--DPNARMTRELKEELAOLRSKIOSGGGGGAGGSGGPVEESYPDOT 413  
 Db 327 akhrintpnhvnevldeallkrykell-----354  
 OY 414 PLEKOIVSIQPOAPT--VKMKSAEIVEQLNQSEKLYRDUNQTMWEKLAETIEHKERPA 472  
 Db 355 dlkkqlenlessetckaqamaee-----htqlaekqlkheredr 396  
 OY 437 LEEGLGSI-----EKGVPVPHSKEMPHLVNLSDDPLLAELCLVYNIKPG 516  
 Db 397 iwhlntivvassgeqgdrvrkrvrcwqqlkslhaagvsdfdmisrl-----448  
 OY 517 QTRVGVNVNDQTQAEIRLNGSKITLKEHCT-FENVNVVTVIPNEKAAYWNGVRDKPRSL 575  
 Db 449 ---pgnfskakkfssdmpsfeldsvctefsdldalmsmds-----ng--lidaewn 496  
 OY 576 RSGYRIILGDPHFIRFNPEEARAEQBSLRHSVTNSQLGSPAPGRHRTLSKAGSDA 635  
 Db 497 as-----kvth-----rektahqsmldfgl-----518  
 OY 636 DQDSRSDSPHFRKGDWFPYARRAASAILGDOKRISHLTDBDELALFDQOKARAVR 695  
 Db 519 -----sdsvqfhdskenkqlgylpkd--sgdmaerckasf--ekeltalqqqlgskeek 569  
 OY 696 KGVEDNE---DSOSGSPFVRKQYMSNTIDNFSLOTATMPTGTPRSDDGDAI-----747  
 Db 570 kelqvsfeklaeieqlsvkxklmvtvnsrsh-sinaev-----qtdvekevrvfrem 622  
 OY 748 -PFGDKSKQSDASNVDFE-----LROQAQAMEAL-----777  
 Db 623 svlgd--sgynaansldqsdsvdkrlseshdeciehrkmlqekivdleefienlnkksae 680

Oy 778 ----KTAKQEF 784  
Db 681 ndkqksseqdf 691  
|:::|

RESULT 9  
ID W72746 standard; Protein; 975 AA.  
AC AC  
XX AC  
XX AC

DT 11-JAN-1999 (first entry)  
DE Drosophila kinesin.

KW Drosophila; kinesin; separation; hybridisation; target site;  
KW complex mixture; motor protein; actively transported; separated;  
KW microtubule.

XX Drosophila sp.

Key Location/Qualifiers  
FH Misc-difference 557  
FT /note= "encoded by AAT"

PN US5830659-A.

PD 03-NOV-1998.

PF 13-SEP-1996; 96US-0713815.

PR 13-SEP-1996; 96US-0713815.

PA (UTAH ) UNIV UTAH RES FOUND.

PI Stewart RJ;

DR WPI; 1998-609236/51.  
DR N-PSDB; V67162.

XX Separation of selected molecules, e.g. DNA, from complex mixtures -  
PT uses specific apparatus to allow the selected molecule to bind to  
PT motor proteins, and be actively transported and separated away along  
PT microtubules

PS Disclosure; Column 17-24; 24pp; English.

XX A method has been developed of separating a selected molecule from a  
CC mixture of molecules. The method comprises: (a) a separation device  
CC comprising a loading reservoir and a receiving reservoir coupled by a  
CC channel with microtubules immobilised on its surface and aligned parallel  
CC to a longitudinal axis of the channel; (b) loading the loading reservoir  
CC with an aqueous solution of the mixture of molecules; (c) adding a motor-  
CC ligand composition and ATP to the solution, where the motor-ligand  
CC comprises, (1) a processive motor capable of attaching to the immobilised  
CC microtubules, and (2) a ligand moving in the presence of ATP as source of chemical  
CC energy, and (3) a ligand coupled to the motor protein, where the ligand  
CC is capable of binding the selected molecule, so that the ligand binds the  
CC selected molecule and the motor protein attaches to the immobilised  
CC microtubules and transports them; and (d) moving the selected molecule from the  
CC receiving reservoir; The method of the system is used for the  
CC separation of specific molecules from complex mixtures. The molecule to  
CC be separated is a DNA ligand. Activation of these enables them to travel  
CC the specific binding ligands. Activation of these enables them to travel  
CC down be performed easily without contaminants of other mixture particles.  
CC The present sequence represents Drosophila kinesin from the present  
CC invention.

XX Sequence 975 AA;

Query Match 16.1%; Score 649.5; DB 19; Length 975;  
Best Local Similarity 27.4%; Pred. No. 5.6e-41;  
Matches 232; Conservative 136; Mismatches 282; Indels 197; Gaps 32;

Oy 6 NIKVVVRVRFENARIDRGAKCIVRMIEGNOTILTPPGAEAKRKSKCTIMDGPKFAFD 65  
Db 12 silkvcrfrpindseekagskfva |:::|  
Oy 66 RSYWSEFDKNAPYARQEDLFODLGVPLLDNAFGYNNICIFATGYGTSGSKSYM---MVG 122  
Db 57 kvf-----kpn-asqekvyeaaksivtdvlagvngtlfayvgqtsgsktmtmegvlds 109  
Oy 123 KEHGVIPIQODMFRRIEQLQDKNLCTVEVSYLEIYNERVROLNPSKTGNKVRHP 182  
Db 110 vkqgiprvindifnhiyam--evalefhikvsvyeiymdkirdldvs-kvnlsvhek 166  
Oy 183 STGPVVEDLAKLVRSFOETIENLMDGKNKARTVAATNMNETSSRSVAFTLTLPOKHDE 242  
Db 167 nrpyvyvgaterfvsspedvfeleedksnrhiavtmnnehsrsrsvflnvkqenlen 226  
Oy 243 ETKMDTEKVAKISLYDLAGSERATSGATGAKLKEGAENRSLSTGLRVIAALADSSCK 302  
Db 227 qkkl-----gklyldvlagsekvstgeetvlddeaknkalsagnvialad---gn 279  
Oy 303 QKNQLVPYRDSVLTMLKDSLGNSMTAMIAISPADINFEETLSTLRYADSAKRKNH 362  
Db 280 kth---ipydsaklrlilqeslgnarttiviccpasfnesekstcltgrfraktvkn 336  
Oy 363 AVNE-----DPNARM---IRELKEELAQLR----- 386  
Db 337 vcvneelLaeewkryekekknarkgkveklelelaragetvkaeeqinmedmea 396  
Oy 387 -----KLQSGGGGGGAG-----GSGGVVEESYPPDTPLEKQIVSIQDPQATVK 430  
Db 397 stpnleveaaqtaaeaaalaagrtalanmsaavvneqarlatecerlyqglldkdeeln 456  
Oy 431 KMSK--AEIVEDQUNSEKLYRDLNQTWEEKLAKTEEIHKEEA-----ALBELGI 478  
Db 457 qsgsqaglkqevmeqeellanaaryetelqsemariqqenesakeevkevlgaeeilcv 516  
Oy 479 SI-----EKFGVGYHSEKEMPHLVLS--DDPLAECLV----- 510  
Db 517 nydqsqeldnknkldalneelqkqsvfnaasteqlqkdmshqkkrilemllnllr 576  
Oy 511 -----YNIKFGQ-----TRGVNWQD--TOAETRLNLSKILKEHCTENGVN 551  
Db 577 dlgevqgalapgeessldkmsalagldskveedfma-rifiskmke-----akul 628  
Oy 552 VTIVPEKAAVWVGVRIIDPTSLSGYRIILGDHIFRNFHPEARPEARQPSLL----- 607  
Db 629 agrcsmetqqadsnkl-----seyekdIgeyrlilishq--eamkslqesmeaen 679  
Oy 608 RHSVYNSQLSGSPAGRHDTLSKAGS-----DADDSRSDPLPHFGKDSQWYARREA 662  
Db 680 krtleedlgl-----reecaklaehvsvnaeekqraeelrsmfsgds-----lrea 732  
Oy 663 ASAILLOKILSHLTD-----EUDALPDVOKARAVRGLVEDNE---DSDSSQSPV 713  
Db 733 ht-----rqvselrdeilaakqhemdemkvnhqkllilhqmqtdykevrqedaekssel 786  
Oy 714 RQXVMSN 720  
Db 787 gniilcn 793

RESULT 10

W72745

ID W72745 standard; protein; 411 AA.

XX AC

XX W72745;

DT 11-JAN-1999 (first entry)



DE Drosophila kinesin N-terminal 411 amino acid residues.  
 XX  
 DE Drosophila; kinesin; separation; hybridisation; target site;  
 KW complex mixture; motor protein; actively transported; separated;  
 KW microtubule.  
 OS  
 OS Drosophila sp.  
 PN US5830659-A.  
 XX  
 XX 03-NOV-1998.  
 PF 13-SEP-1996; 96US-0713815.  
 XX  
 PR 13-SEP-1996; 96US-0713815.  
 XX  
 PA (UTAH ) UNIV UTAH RES FOUND.  
 XX  
 PI Stewart RJ;  
 XX  
 DR WPI; 1998-609236/51.  
 XX  
 XX Separation of selected molecules, e.g. DNA, from complex mixtures -  
 PT uses specific apparatus to allow the selected molecule to bind to  
 PT motor proteins, and be actively transported and separated away along  
 PT microtubules  
 XX  
 XX Claim 3; Column 25-28; 24pp; English.  
 XX  
 XX A method has been developed of separating a selected molecule from a  
 CC mixture of molecules. The method comprises: (a) a separation device  
 CC comprising a loading reservoir and a receiving reservoir coupled by a  
 CC channel with microtubules immobilised on its surface and aligned parallel  
 CC to a longitudinal axis of the channel; (b) loading the reservoir  
 CC with an aqueous solution of the mixture of molecules; (c) adding a motor-  
 CC ligand composition and ATP to the solution, where the motor-ligand  
 CC comprises, (i) a processive motor capable of attaching to the immobilised  
 CC microtubules, and (ii) a ligand coupled to the motor protein, where the ligand  
 CC is capable of binding the selected molecule, so that the ligand binds the  
 CC selected molecule and the motor protein attaches to the immobilised  
 CC microtubules and transports the bound selected molecules along the  
 CC receiving reservoir; and (d) removing the selected molecules from the  
 CC separation device.  
 CC Separation of specific molecules from complex mixtures. The molecule to  
 CC be separated (e.g. DNA) binds to the motor protein. The presence of  
 CC the specific binding ligands. Activation of these enables them to travel  
 CC down a preformed channel in a specially made piece of apparatus. They can  
 CC then be removed easily without contaminants of other mixture particles.  
 CC The present sequence represents the N-terminal 411 amino acid residues of  
 CC Drosophila kinesin for use in the method of the invention.  
 XX  
 SQ Sequence 411 AA;

Query Match 15.5%; Score 625; DB 19; Length 411;  
 Best Local Similarity 39.4%; Pred. No. 1.1e-39;  
 Matches 134; Conservative 67; Mismatches 128; Indels 42; Gaps 11;  
 OY 6 NIKYVVRVFNAHEDINDGACIKTYVMGNOQILTPPGAEAKRSKGTIMDGFPAFAD 65  
 DB 12 SLKVCRTFDIndeekagskfvrk-----pnvnevc-----isiagkyvld 56  
 OY 66 RSTWSFDKNAHVARQEDFDGLGVPLIDNAFGYNNCTFEAYGOTGSGKYSVM--NWGT 122  
 DB 57 kvf-----kpn-esqekvynaeaaisvcdvlagyngtifaygtsqgkhtmevgld 109  
 OY 123 KENGHVPICQDFRINELQDKMLTCVTEVSLYIENVRDLNWPSTKMLKAVRHP 182  
 DB 110 vkqgflpirdvndfnliyam--evnlefkivsvyelymdkirdldvs-kvnlsvhedk 166  
 OY 183 STGPYVEDAKLVRSFOETIMDECNKARTVAATNNNETSSRSHAVPTLTITQKHDE 242

DB 167 nrvyvgkaterfsspedvfeieegksnrhiavtumehsrsahvflnvkqenlen 226  
 OY 243 EHKMDTEKVAKISLVDLAGSERATSTGATGARLKEGATNRSLSLGTGVTAALADMSGK 302  
 DB 227 qkks-----gklylvldagsekvsktgdegtvldeaknlnksalgnvisalad---gn 279  
 OY 303 QKKNLVPPYRUSVLTWLLKDSIGNSMTAMIAISPADINFEETLSLTARYADSAKRIKH 362  
 DB 280 kch---lpyrdskltrilgdeslgnartiviccspasfneseetkstdlfdgraktvknv 336  
 OY 363 AVWEDPNA----RMIRELKRELAQLRSKIQ 389  
 DB 337 vcvneelaaewkrrykekenarikgkve 367  
 RESULT 11  
 W72744  
 ID W72744 standard; protein; 441 AA.  
 AC W72744;  
 XX  
 XX 11-JAN-1999 (first entry)  
 DT  
 XX Drosophila kinesin N-terminal 441 amino acid residues.  
 DE  
 XX Drosophila; kinesin; separation; hybridisation; target site;  
 KW complex mixture; motor protein; actively transported; separated;  
 KW microtubule.  
 OS  
 OS Drosophila sp.  
 PN US5830659-A.  
 XX  
 XX 03-NOV-1998.  
 PF 13-SEP-1996; 96US-0713815.  
 XX  
 PR 13-SEP-1996; 96US-0713815.  
 XX  
 PA (UTAH ) UNIV UTAH RES FOUND.  
 XX  
 PI Stewart RJ;  
 XX  
 DR WPI; 1998-609236/51.  
 XX  
 XX Separation of selected molecules, e.g. DNA, from complex mixtures -  
 PT uses specific apparatus to allow the selected molecule to bind to  
 PT motor proteins, and be actively transported and separated away along  
 PT microtubules  
 XX  
 XX Claim 3; Column 23-26; 24pp; English.

A method has been developed of separating a selected molecule from a mixture of molecules. The method comprises: (a) a separation device comprising a loading reservoir and a receiving reservoir coupled by a channel with microtubules immobilised on its surface and aligned parallel to a longitudinal axis of the channel; (b) loading the loading reservoir with an aqueous solution of the mixture of molecules; (c) adding a motor-ligand composition and ATP to the solution, where the motor-ligand comprises, (i) a processive motor capable of attaching to the immobilised microtubules and moving in the presence of ATP as source of chemical energy, and (ii) a ligand coupled to the motor protein, where the ligand is capable of binding the selected molecule, so that the ligand binds the selected molecule and the motor protein attaches to the immobilised microtubules and transports the bound selected molecules along the receiving reservoir; and (d) removing the selected molecules from the separation device.  
 Separation of specific molecules from complex mixtures. The molecule to be separated (e.g. DNA) binds to the motor protein due to the presence of the specific binding ligands. Activation of these enables them to travel down a preformed channel in a specially made piece of apparatus. They can then be removed easily without contaminants of other mixture particles.



PR	14-OCT-1999;	9905-0159637.	Query Match	14.58;	Score 584;	DB 21;	Length 1518;
PR	14-OCT-1999;	9905-0159638.	Best Local Similarity	29.38;	Pred. No. 1.2e-35;		
PR	18-OCT-1999;	9905-0159584.	Matches 196;	Conservative 102;	Mismatches 195;	Indels 176;	Gaps 27;
PR	21-OCT-1999;	9905-0160767.					
PR	21-OCT-1999;	9905-0160768.					
PR	21-OCT-1999;	9905-0160770.					
PR	21-OCT-1999;	9905-0160814.					
PR	22-OCT-1999;	9905-0160815.					
PR	22-OCT-1999;	9905-0160980.					
PR	22-OCT-1999;	9905-0160981.					
PR	25-OCT-1999;	9905-0161404.					
PR	25-OCT-1999;	9905-0161405.					
PR	25-OCT-1999;	9905-0161406.					
PR	26-OCT-1999;	9905-0161359.					
PR	26-OCT-1999;	9905-0161361.					
PR	28-OCT-1999;	9905-0161920.					
PR	28-OCT-1999;	9905-0161992.					
PR	28-OCT-1999;	9905-0161993.					
PR	29-OCT-1999;	9905-0162142.					
Qy	17 NAREIDRGAKC	---IYVWBGNGQTILTPPPGAEBKARKSKGKTIMDGKAFAPDRYSWSD 72					
Db	41 nappdrntspdh:simnk-nlplppppsnplpkraytsaetates-----glds----- 90						
Qy	73 KNAPNVAQRQEDLPDGLVGLVQVGGKAGVYNNCTAYFGGSGSGSKYSWAG-----YG 122						
Db	91 -----gvkeqmqlqvgaaplvcnclsgfnssvfyggtgqctymvgpanglleehlq 144						
Qy	123 KEHGVIPRIQDQMFRRINE-----LQDKNNTTCTVEVSYLEYIYNEVRDLIN- 169						
Db	145 dgrqlprvrlfarlkevmgfvqvitnksmlkgn-sitcnvgeghysrfadwlhy 203						
Qy	170 -----PSTKGNL-----KYREHPSTGPTVEDLAKLVRSFDE 201						
Db	204 lsldlqranrptpkpkkpdkghvhrfslmchqlrdevggyvnenlceeyvknld 263						
Qy	202 TENLDGNGKARVAVAAWNTSSSHAVPTLITQKWHDETKMDTFEYAKVLSVLQAG 261						
Db	264 vsqllkglgnrttgatsvntssrshovftcvcearcknvadgljaefktarinldvg 323						
Qy	262 SERATSGTGAARLKEGAENRSLSTLGRVIAALADMSSGKKKNOLVPRVRSVLWLL 320						
Db	324 serqctgaqetlkeaglnrslsglnlnlslaegtkprh---lpydrslrflil 380						
Qy	321 KDSLGGNSMTAMTAISAIPFTEETSLTRYADSAKRKNIAVYNN-----DPN--ARMI 374						
Db	381 qeslgnaklamvcvsepsqrs-etfstlrfagrkaglnkavnmvqmdvfnlrgvl 438						
Qy	375 RELKEEAQLRSL-----LQSGGGGGGAGSGGGVPEESYPP 411						
Db	439 hqldrlqkndgndgnptnpyastawarrslnllstlrgly-----lpr 484						
Qy	412 DTPLEKXO--TVSIQQDPATVKKSKAEIYFQNLQNSKLYLNDLNOTWEEKLAKTEIHKER 469						
Db	485 slphedngldemeldeaaverlcvqvgqlqsslsseglnhdmm-----rvkslhead 536						
Qy	470 EAALEGLSIEKGFGVGVSHKEMPHLVNLSDPLLAELCLVYNIKPGQTVGVGNWQDTQA 529						
Db	537 -----ggslkrlr--pedad-----vamed-----acchtenhep-etvdmrmetet 576						
Qy	530 ETRLN-----GSKLKEHCTF-----ENVDVVITYVPEKAA-----VMVNG 566						
Db	577 girenglktlnsqldchessfqlsvkdalosslnkdsedvscpdlpqvgvdsenylldg 636						
Qy	567 VPIKDPZBL 575						

Db 637 v--ddpehl 643

RESULT 13

G31282

ID G31282 standard; Protein; 1662 AA.

XX G31282;

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 37541.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

XX Arabidopsis thaliana.

OS

XX

PN EF1033405-A2.

XX

PD 06-SEP-2000.

XX

XX 25-FEB-2000; 2000EP-0301439.

PF

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 06-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134236.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134231.

PR 14-MAY-1999; 99US-0134232.

PR 14-MAY-1999; 99US-0134233.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145921.

PR 28-JUL-1999; 99US-0145921.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

[illegible]

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138450.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141281.  
PR 01-JUL-1999; 99US-0142152.  
PR 01-JUL-1999; 99US-0142055.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142303.  
PR 08-JUL-1999; 99US-0142303.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145142.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0152139.  
PR 23-SEP-1999; 99US-0152456.  
PR 28-SEP-1999; 99US-0152458.  
PR 28-SEP-1999; 99US-0156498.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157665.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158332.  
PR 08-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.

PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 14.4%; Score 580.5; DB 21; Length 1460;			
Best Local Similarity 29.4%; Pred. No. 2e-35;			
Matches 189; Conservative 99; Mismatches 183; Indels 171; Gaps 25;			
QY	40	PPGAEAKRSKGTMTGPKAFDRSFWSPDKNAPYARQEDLFDQLGVPLLDNAFKG 99	
DB	9	pppsnpIkrKlaetates-----gfads-----gvkeqmfqlvgapIvencIsq 53	
QY	100	YNICFIAYGOTGSGKSYSNMG-----YKHEGVIPIRCODMFRINE----- 141	
DB	54	fnsfvfyqctsgagkytcmgpanllleehLcgdrgltprvfierIarIekwmpgfvqf 113	
QY	142	----LQDKMLCTVENSYLEIYNERVRLAN-----PSTKGNL 176	
DB	114	vvItInksmlkgn-sItngabhyrsfdrwIhYIsIdIdIqrannrtgpkpdkgkhg 172	
QY	177	-----KYREHSPGYPVEDIAKLVRSFOETENLMDGKNKARTVAATNNHETSSRH 228	
DB	173	hfrsImchqIredvkagvyvenIteeyvknItdvsqllkIglgrrtgatsvntesrsh 232	
QY	229	AVFTLTITOKWHEETKMDTEKVKARISLVDLAGSERATSTGATGARKEAGAEINRSLSLT 288	
DB	233	cvfIcvesrcnkvadgIassfktrInIvdlagserqstgaagerIkeagnInrsIsqL 292	
QY	289	GRVIAALADMS--SKOKKQNLQPYRDSVITMLKSLDGGNSMTAAIASPADINFEETL 347	
DB	293	gnInIlaeIsctqkphr---IpydrsrItIqlsagIknIamvcavpsqr--etf 347	
QY	348	STLRVADSARKTKHAYNE-----DPN--AKWIRELKEELARSK----- 387	
DB	348	stIrfaqrakIqnkavnermqdavnfIrgIhqlrdeIqkmdgnpntpuyavsta 407	
QY	388	-----LOSSGGGGGAGSGGCPVEESYPDPLEKQ--TVSIOQPDATYKVKSKAE 436	
DB	408	wnarrsInallrfsqI-----hpsIphedngdIemeIdeaaVerlcvg 453	
QY	437	IVPQNLNSEKLYRLNQTWEKIATBEIHKREAAALBELGISTEKGFVGYHSKEMPHL 496	
DB	454	gIqslasegInhdnn-----rvksIhasd-----ggsIekrl--peds----- 491	
QY	497	VNLSDDPLAELCYNIKPGOTRVGNVNOOTQAEIRLN-----GSKILKEHCTP----- 545	
DB	492	vamed-----acchenhep---etvdmmrtetqIrenqIkthqtdIhessqIsvkd 545	
QY	546	-----BNVDNVTIVPNKAA---WVNGVRIDKPRPEL 575	
DB	546	alcsalnksdsvscpdlvpqdrtsanvIadgv--ddpehl 585	

RESULT	15
GA0076	G40076 standard; Protein; 1462 AA.
XX	G40076;
XX	18-OCT-2000 (first entry)
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 49677.
XX	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	Arabidopsis thaliana.
XX	EP1033405-A2.
XX	06-SEP-2000.
XX	25-FEB-2000; 2000EP-0301439.
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130681.
PR	28-APR-1999; 99US-0130749.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	18-MAY-1999; 99US-0134370.
PR	19-MAY-1999; 99US-0134376.
PR	20-MAY-1999; 99US-0134941.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	27-MAY-1999; 99US-0136041.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	14-JUN-1999; 99US-0139119.
PR	16-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	17-JUN-1999; 99US-0139492.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.





Search completed: April 25, 2001, 10:12:41  
Job time: 337 sec

**THIS PAGE BLANK (USPTO)**



Db 291 -KKRSDFIPYRDSVLTWLLKENLGNRSRTAMTAAALSPADINVEETLSTLRVADRTKQIR 349  
 QY 361 NIAVVDNPNARIRELKEELADLQSLKSGGGGGGAGG----- 400  
 Db 350 CNAIINEDPNARIRELQEEVARSRELLMAQGISASALEGLKTEBSVRGALPAVSSPPA 409  
 QY 401 -----SGPVESYPPDPLQELQVSIQDPATVKKMSKAEIVQDLNQSSEKLYRL 451  
 Db 410 PVPSSPTTHNGELPESFSPNT--ESQI-----GPEAMERLQETEKIAEL 454  
 QY 452 NOTWEEKLAKTEIHKEREALDELGISIEK--GFVGPYHSEKMPHLVNLSDPLLAEL 509  
 Db 455 NETWEEKLRKTEALREMERALLAEMGVAVREDGTVGVFSKPTPHLVNLNEDPLNSECL 514  
 QY 510 VYNIKPGOTRGVNVNDQTQAEIRLNGSKILKEHCTFENV-----DNVVTIVPNEKAAMV 564  
 Db 515 LYHFKGVTRGVGDMD-----IKLTGQITREQHCLFRSIPQPDGEVVVTLPCGEGATTV 570  
 QY 565 NGVRIDKPTRLRSGYRIILGDPHFRFNHPFEARAEQESLLRHSVTSQILSGSPAPGRH 624  
 Db 571 NGKLVTEPLVLKSGNRYVHGKNHFRFNHPFEQARLER----- 614  
 QY 625 DRTLSKAGSDADGSDSDSPHIFRFGKDSQDFVARRRASAIGLDQKISHLTDDELAL 684  
 Db 615 -----PGPFPSEPDVWNEFAQKELLEQ--QGIDIKLE--MEKRLQDL 650  
 QY 685 FDDVQKARVRRLGVEDNE---DSOS 707  
 Db 651 ENQYRKEEEADLLLEQORLYADSDS 676

## RESULT 2

US-08-006-676B-1  
 ; Sequence 1, Application US/08006676B  
 ; Patent No. 5411865  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven  
 ; TITLE OF INVENTION: Diagnosis of Leishmaniasis  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Jeffrey B. Oster  
 ; STREET: 8339 SE 57th Street  
 ; CITY: Mercer Island  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98040-4906  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WORD for Windows  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/006,676B  
 ; PUBLICATION: 15 JAN-1993  
 ; CLASSIFICATION: J45  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Oster, Jeffrey B.  
 ; REGISTRATION NUMBER: 37,595  
 ; REFERENCE/DOCKET NUMBER: REED-4  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (206) 232 7845  
 ; TELEFAX: (206) 235 0205  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 955 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-006-676B-1

Query Match 20.2%; Score 814; Db 1; Length 955;  
 Best Local Similarity 28.1%; Pred. No. 7,4e-60;

Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;  
 QY 7 LKVVVRVPPFNARE--IDRGAKCIVRMGGNOTILTPP-----PGAEEKARKSGKTI MD 57  
 Db 13 YKVSVRVPPPLNERENNAPEGTKYVAAQAAAVTVTVKVGSGNNSGAESMGTARRVAQD 72  
 QY 58 GPKAFKASVWYGF--DKNAPNARQEDLPDGLGVPLLDNFAKYNKICIFAYQOTGSGK 114  
 Db 73 ----FQDHVFWSVETPDACGATPATQADVFTICGYPLVQHAFDGFNSCLFAYQOTGSGK 128  
 QY 115 SVSNWNG-----YGKEHGVIPRICODMPFRINLQKDNLTCTVSEVSYLEIYNERNVDLL 168  
 Db 129 TVTKMGADYALSQGGNGVTPRICLEIFAPKASVEAQGSRMIVELCYVEVTVNERVSDLL 188  
 QY 169 NPSYKQ-----NLKYRHPSTGPYVEDLAKLVRSFOEINEMDGNKARTVAATNN 220  
 Db 189 KRRKKVGGGGEVYVDVREHPSRGVFLQGRVLVWVGLSDVDVRLIEIGNVHRTASTKM 248  
 QY 221 NETYSRSHAVFTLT-----TOKWHEDET KMDTEKVAKISLVLDAGSRATSTGATGARKL 276  
 Db 249 NDRSRSHAI TMLLREERTWTKGETIRTAGKSRMNLVLGSRVAQSVQSGQGF 308  
 QY 277 EGABINRSLSTGRVFAALADMS--GKKKNQLVIPYRDSVLTWLLKSLGGSNTAMIAA 335  
 Db 309 EATHINLSLTGRVIDVLADMTAKAKAQSVPAPFRDSKLTFLKSLGSGSKTFMIAT 368  
 QY 336 ISPADINFEETLSTLRYSADSAKRKNHAWNEDPNARMRELKEELAQRLKSLQSGGG 395  
 Db 369 VSPSALNVEETLSTLRYSARQIVNVAOVNEDPRARRIRELEEQMEDMFOAM----- 421  
 QY 396 GGAGSGGGPVEE-----SYPPDTPLEK 417  
 Db 422 --AGDPATVYSELKKLALLEBAQKRAADLQALEREHNOVQERLLRATAEKSELES 479  
 QY 418 QIVSTIQQ-----PDATVKMSKAEIYEOL 441  
 Db 480 RAAALQEEMTATRRQADKMQALNLRKEEQARKERELKEMAKKDAALSXVRRKDAETA 539  
 QY 442 NOSEKLYRLDNLQWTEELAKTERIHKEREALBELGISIEKGFVGYHSEKMPHLVNLSD 501  
 Db 540 SERELK-----ESTVAOLEREOREVALDAL-----OTHOKLOEALLESSE 581  
 QY 502 -----DPLLAELVYNIKPGOTRGVNVNDQTQAEIRLNGSKILKEHCTFENVDNVYTV 555  
 Db 582 RTAAERDQLAQ--LTLEQSERTQLSQVVDRELRTR-DLQRTQYEGETELARDVALCA 638  
 QY 556 PNE-----KAAVWVGVTRDKPTRLRSGYRIILGDPHFRFNHPFEARAEQESLLRHS 610  
 Db 639 AQEMERYHAHVFILOTLELATMEDALR-----BRALAEDEAAAAE-- 682  
 QY 611 VTNQSGSPAPGRHIDRTLSKAGSDAGDSRSDSPHPFRCKSDWFTARRAASAIGLD 670  
 Db 683 -----IDNAASTSQN-----ARESCRELSTLE 705  
 QY 671 QKISHLTDDLDALFDVQKARAVRGLVEDNEDSDSSSFVPKRYKTSNGTINDFSITD 730  
 Db 706 QQLRE--SEERAELASQLEATAAKSSAQEDRENTATLLEQURES-----FARAEL 757  
 QY 731 ATTPGTPRSDODGDALFFGDKKSKODASNV-----DYRELROGQOMEALAKTAK 781  
 Db 758 ASQLEATAA-----KMSAQEDRENTATLLEQLEDSEERAAELASOLESTTAAK 807

## RESULT 3

US-08-282-845-2  
 ; Sequence 2, Application US/08282845  
 ; Patent No. 5719263  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; TITLE OF INVENTION: A 230kd Antigen Present in Leishmania  
 ; NUMBER OF SEQUENCES: 3  
 ; CORRESPONDENCE ADDRESS:

ADDRESSSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Operating System 7.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282,845  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/006,676  
FILING DATE: JANUARY 15, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 5004-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 955 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-282-845-2

Query Match 20.2%; Score 814; DB 1; Length 955;  
Best Local Similarity 28.1%; Pred. No. 7,4e-60;  
Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;  
QY 7 IKVVVVRFPNARE--IDRGAKCIVRMGNGQITLTP-----PGAEEKARKSKGTMID 57  
DB 13 VNSVVRPLNERENMAPEGTKYVAAQAAVYVYKGGNSNGMESMTARRVAQD 72  
QY 58 GPRAFAFDRTSYSP---DKNAPHYARQEDFDGLGVPLLDNAPKYNKCIPTNGTQSGK 114  
DB 73 ----FQDRHVSFETPDACGATPATQADVPRITGYPLVQHAFGNSCLFALPGTQSGK 128  
QY 115 SVSMG-----YKKEHGVPICDMFRNELQNDKNTCTVSELYETNEVRDL 168  
DB 129 TYTHMGADVSALEGGNGVTRICLEIFARKASVEAQGHSMVELGYVEYNERVSDLL 188  
QY 169 NFSTG-----NLKVREHPSGTGVVEDAKIVIRFSOEIENLMDENKARTVAATM 220  
DB 189 GKRRKGVKGGGEYVDVREHPSRGVLEGQRLVEVGSLODVYRLTEIGNVGHRTAKM 248  
QY 221 NETSSSHAVFTL-----TQKHDEPTKMDTEKVAKISLVLAGSPRATSTGATGARKL 276  
DB 247 NDHSSKSHATIMILLREERTMTYKGGETIRTAGKSSRNMLVLAGSERVAQSGVGGQFK 308  
QY 279 EGAEINSLSTLGRVITALLADNGS--GKQKKQLVPRVSVLTWLLKDSLGNSMTAMIAA 335  
DB 309 EATHINSLTIGRVIDVLAADKATGAKAQSVNAPRMSKLFTFLKDSLGNSKTFMIAT 368  
QY 336 LSPADINPEETLSLTAYADSKATVHVNVEDPNKMRBELKELAQRLSKLSSGGCG 395  
DB 369 VSPSALNYEETLSLTAYASRARDIVNVAQVEDPRARELRELEQMEDMDRQAM----- 421  
QY 396 GGAGSGGPGVE-----SYPPDPLEK 417  
DB 422 --AGDPAPVYSELKKLALLESPAQRADLQALERREHNOVQERLLRATAEKSELES 479  
QY 418 QIVSIQO-----PDATVKKMSKASITVEOL 441

RESULT 4  
PCT-US94-00324-1  
Sequence 1, Application PC/TUS9400324  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven  
TITLE OF INVENTION: Diagnosis of Leishmaniasis  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple System 7.1  
SOFTWARE: Microsoft Word, version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/00324  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/006,676  
FILING DATE: 15-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 5004-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 955 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-00324-1

Query Match 20.2%; Score 814; DB 5; Length 955;  
Best Local Similarity 28.1%; Pred. No. 7,4e-60;  
Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

QY 7 IKVVVRVPFNARE--IDRGAKCIVRMEGNOTILTPP-----PGAEKARKSGKTMID 57  
DB 13 VKVSVRVPFNARENNAPGECTKVYAAQAAAVVTVKLVGGSNNSGAESMGTARRVAOD 72  
QY 58 GKAFADRSWFSF--DKNAPNARQEDLFDGLVPLDPAFNGYNNCIFAYGOTGSGK 114  
DB 73 ----FOFBHFWSVETPPDAGCATPATQADVFTICGYLVGHAFDGFNSCLFAYGOTGSGK 128  
QY 115 SYSMMG-----YGRHEGVIPRICODMFRINELQKKNITCTVEVSYLEIYNBRVDLL 168  
DB 129 TYTMKADVSAISGEGNVTPRICLEIFPARKASVEAOGHSRWIVELGYVEVYNERVSDLL 188  
QY 169 NPSTKG-----NLYRHPSTGPVVEDAKLVRSFOETENIMDEGNKARVATNM 220  
DB 189 GKRRKVGKGGEVYVDRHPSRGVFLQGRVLEQGLVDVVRLLIENGVRITASTKM 248  
QY 221 NETSSSHAVFTLL-----TKQWHEETKMDTEKVAKISLVLAGSERATSTGATGARLK 276  
DB 249 NDRSSSHAIIMLLREERTWTTKSETITAGKSSRMNIVDLAGSRVAQSQVEGQCFK 308  
QY 277 EGAEINRSLTGRIATADMS--GKOKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAA 335  
DB 309 EATHINLSLTGRIVDLADMATKGAQYVAPFRDCKLTFILKDSLGGNSKTFMIAT 368  
QY 336 ISPADINFEETSLTRYADSAKINKHAYNEDPNARMIRELKEELAQRLSKLQSSGGG 395  
DB 369 VSPSALNTEETSLTRYASARDIVNVAQVNEEDPRARRIRELEBQMEDMQAM----- 421  
QY 396 GKAGSGGPGVEE-----SYPPDTPLEK 417  
DB 422 --AGDPAYVSELKXKLLALESPAQRAADLQALREBRHNQVORLLRATEAEKSELES 479  
QY 418 QIVSIOQ-----PDATVKMSKAEIVFOL 441  
DB 480 RAALAEQEMTATROADKMOALNLRKEQARKERELKEMAKDAALSKVRRKDAETA 539  
QY 442 NOSKLYHDLNQTWEKLANTEITHKREALEELGISTEKGFGVPHSKEMPHLVNLS 501  
DB 540 SERKJ-----ESTVAQLEREQOREVALDAL-----QHQKRIQEALESSE 581  
QY 502 -----DPLAELVYNTKQTRGVNVDQTAERLNGSKILKEHCTFENVNDVTV 555  
DB 582 RTAERDQLQO--LELQSERQTQSQVYDRERLTR-DIQRITQVETGETELARDVALCA 638  
QY 556 PNE-----KAQVWNCVRIDKPTRLSQYRIILGDPHIFRPHNPEARAEQOSLIHRH 610  
DB 639 AQEMEARVYAAVPHLOTLELATWEDALR-----ERLAERDEARAAE-- 682  
QY 611 VTNSQLSGPAGRHDTLSKAGSDAGDSUSPLPHFRKQSDWFFARREANAAILGD 670  
DB 683 -----LDAASTSQN-----AESNCERLISLE 705  
QY 671 OKLSHTLDELDLFDVQKARVGLVEDNEDSDQSFPVRDKYMGNTINFSLDT 730  
DB 706 QOLRE--SEERAELASOLEATAAKSSNAQDRENTATRYLEQJRES-----PARAEL 757  
QY 731 AITWPGTPRSDDDGDALPFGDKKSDASVY-----DVEELROQAOEWELKATK 781  
DB 758 ASOLEATAA-----KMSAQDRENTATRYLEQJRESERAAELASOLESTTAK 807

RESULT 5  
US-08-428-414A-3  
; Sequence 3, Application US/08428414A  
; Patent No. 5912166  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; TITLE OF INVENTION: LEISHMANIASIS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY  
; STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428.414A  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Kadlecak, Ann T.  
REGISTRATION NUMBER: 39,244  
REFERENCE/DOCKET NUMBER: 210121.407  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEEDANDBERRY  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 955 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
US-08-428-414A-3

Query Match 20.2%; Score 812.5; DB 2; Length 955;  
Best Local Similarity 40.0%; Pred. No. 9.9e-60;  
Matches 201; Conservative 75; Mismatches 164; Indels 63; Gaps 13;

QY 7 IKVVVRVPFNARE--IDRGAKCIVRMEGNOTILTPP-----PGAEKARKSGKTMID 57  
DB 13 VKVSVRVPFNARENNAPGECTKVYAAQAAAVVTVKLVGGSNNSGAESMGTARRVAOD 72  
QY 58 GKAFADRSWFSF--DKNAPNARQEDLFDGLVPLDPAFNGYNNCIFAYGOTGSGK 114  
DB 73 ----FOFBHFWSVETPPDAGCATPATQADVFTICGYLVGHAFDGFNSCLFAYGOTGSGK 128  
QY 115 SYSMMG-----YGRHEGVIPRICODMFRINELQKKNITCTVEVSYLEIYNBRVDLL 168  
DB 129 TYTMKADVSAISGEGNVTPRICLEIFPARKASVEAOGHSRWIVELGYVEVYNERVSDLL 188  
QY 169 NPSTKG-----NLYRHPSTGPVVEDAKLVRSFOETENIMDEGNKARVATNM 220  
DB 189 GKRRKVGKGGEVYVDRHPSRGVFLQGRVLEQGLVDVVRLLIENGVRITASTKM 248  
QY 221 NETSSSHAVFTLL-----TKQWHEETKMDTEKVAKISLVLAGSERATSTGATGARLK 276  
DB 249 NDRSSSHAIIMLLREERTWTTKSETITAGKSSRMNIVDLAGSRVAQSQVEGQCFK 308  
QY 277 EGAEINRSLTGRIATADMS--GKOKNQLVPYRDSVLTWLLKDSLGGNSMTAMIAA 335  
DB 309 EATHINLSLTGRIVDLADMATKGAQYVAPFRDCKLTFILKDSLGGNSKTFMIAT 368  
QY 336 ISPADINFEETSLTRYADSAKINKHAYNEDPNARMIRELKEELAQRLSKLQSSGGG 395  
DB 369 VSPSALNTEETSLTRYASARDIVNVAQVNEEDPRARRIRELEBQMEDMQAM----- 421  
QY 396 GKAGSGGPGVEESYPPDTPLEKQVLSYQQPDATVKMSKAEIVELQNSKEKLYRDLNQTW 455  
DB 422 --AGDPAYVSE-----LKKKLLIES-----BAQKRAADLQALERE-----REHNOVO 463  
QY 456 EKILAKTEITHKE--REAALEE 475  
DB 464 ERLLRATEAEKSELESRAALQOE 486

RESULT 6  
US-08-713-815A-4



Sequence 48, Application US/08545860D  
Patent No. 6040140  
GENERAL INFORMATION:  
APPLICANT: Croce, Carlo  
APPLICANT: Canaan, Eli  
TITLE OF INVENTION: Diagnostics, Therapeutics and Methods  
TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias  
TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &  
ADDRESS: No. 6040140ris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 07-MAR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/04496  
FILING DATE: 22-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/10930  
FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/327,392  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/320,559  
FILING DATE: 11-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/062,443  
FILING DATE: 14-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/971,094  
FILING DATE: 30-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,839  
FILING DATE: 27-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/805,093  
FILING DATE: 11-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca Esq., Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1262  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3100  
INFORMATION: 48  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1612 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

Query Match 4.88; Score 195; DB 3; Length 1612;  
Best Local Similarity 20.4%; Pred. No. 2.4e-07;  
Matches 139; Conservative 90; Mismatches 241; Indels 210; Gaps 28;

QY 43 GAEKARKSGKTIMDPKAFADRSYNSFOKNAPYAROEEDLPDOLGVPLLDNAPKCYNN 102  
DB 5 GRDEERKLADI1-----HWN-----ANRLDLP-EISOPTEDLEFHGVR 44  
QY 103 CIFAQGTG--SKSYSWVGREGVIPRICODMFRINELOKDKNLACTVEVSYLEY 160  
DB 45 FVFDKAAAGNFATKCI RVSSATTQDTEILAE-----KRPDMRLSSPKSYLYEVH 97  
QY 161 --NERYVDLLNPSTKGNLKVREHSPSTGPGYVEDLAKLVVRSFQETENLMDGSKARTVAAT 218  
DB 98 VSGERRLDIDEKPLVQLNNKNDREGFV-----LKNENDAIPPKQAS 141  
QY 219 NNNETSSSHAV--FTLTQKWHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARLK 276  
DB 142 NGPEQKEGVIONEKRTLSKKEKKREKALRQASDKO-----DRPF 187  
QY 277 EGAEINRSLSTLGRVIAALADMSSGKOKKQNLVPRVDSVLTWLLKDSLGSNSMTAMIAAI 336  
DB 188 QGEDVENS-----RLAEV-----YKDM----- 205  
QY 337 SPADINFEETLSRYADSAKRIKHAVNEDFNARKIRELKEPEALRSKLQSSGGGG 396  
DB 206 --PETSFTRTIS-----NPEVVKRRROOKLEKRQOEFRSS-----D 240  
QY 397 GAGSGGPVE--ESYPDPPTPLEKQIVSIQOPDAIVKMKSAEIVQOLNQSKLYRDL-- 451  
DB 241 CRPDSGGTLRIYADSLKPNIPYKTILLSTTDP-----ADPAVAEALKEYGLEKNPKDYCI 296  
QY 452 -----NOTWEEKLAKTEIEHKREAALE-----ELGSI----- 480  
DB 297 ARVMLPPGAQHSDEKGAKEIILDDDECPILQIPREWPDSKGLVFLQKRRPDPHPKTK 355  
QY 481 -----EKGFVGPYHS-----KEMPHLVNLS-----DDPLAECLVYNIKFGOT 518  
DB 356 KHEGKTPKGRKADSGVYGTSLPPEKLPVLVELSPGSDSDSRDKPKL-----YRIQLSVT 410  
QY 519 RVGNVQDTPQAEIRLNGSKILKECHTFENVNDVYTVFNE-KAAYMVNGVRIDKPTLRAS 577  
DB 411 EVGTEKLDNNS-IQLFGGQIPHCDLTNMDGVTVTPRSMOAEYTYVEGORISSETMLOS 469  
QY 578 GYRIILGDFHIFRNHPPEARAEROESLLRHSVTNSQLGSPAGRGHRTLSKAGSDAG 637  
DB 470 GMYVFGASHVFKFVDPDS-----QDHALAKRSV-DGGLMYGPRKPGIVQETTFDLOG 522  
QY 638 DSRSDSPLPHFRGK---DSD 654  
DB 523 DIHSGTALPSTKSTRLDSD 542  
RESULT 9  
PCT-US94-04496-48  
Sequence 48, Application PC/YUS9404496  
GENERAL INFORMATION:  
APPLICANT: Croce, Carlo  
APPLICANT: Canaan, Eli  
TITLE OF INVENTION: Diagnostics, Therapeutics and Methods  
TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias  
TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &  
ADDRESS: No. 6040140ris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:



```

RESULT 10
US-09-157-420-1
: Sequence 1, Application US/09157420
: Patent No. 6180760
: GENERAL INFORMATION:
: APPLICANT: TAKAI, Yoshimi
: APPLICANT: KANANISHI, Hiroyuki
: APPLICANT: NANDAI, Kenji
: APPLICANT: WADA, Manabu
: APPLICANT: OBARASHI, Hiroshi
: TITLE OF INVENTION: ACTIN FILAMENT-BINDING PROTEIN "L-AFADIN"
: FILE REFERENCE: 98-1042-JC(WMC)/653
: CURRENT APPLICATION NUMBER: US/09/157,420
: CURRENT FILING DATE: 1998-09-21
: CURRENT SEQ ID NOS: 1
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO. 1
: LENGTH: 1829
: TYPE: PPT
: ORGANISM: rat
US-09-157-420-1

```

Query Match 4.6%; Score 185.5; DB 4; Length 1829;  
Best Local Similarity 21.1%; Pred. NO. 1.8e-06;  
Matches 148; Conservative 91; Mismatches 265; Indels 199; Gaps 32;

[illegible]

QY 638 DSRSDSPLPFRGK---DSDF-----YARREASAILGLDOK 672  
 DB 539 DIHSGTALPASRSTRLSDRVSASSSTABRGWKPMLRDOE 581  
 RESULT 11  
 US-08-353-700-1  
 ; Sequence 1, Application US/08353700  
 ; Patent No. 5599919  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YEN, TIMOTHY J.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A  
 ; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,  
 ; TITLE OF INVENTION: AND METHODS OF USE  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN  
 ; STREET: 1601 MARKET STREET, SUITE 720  
 ; CITY: PHILADELPHIA  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19103-2307  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; FILING DATE: 09-DEC-1994  
 ; APPLICATION NUMBER: US/08/353,700  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: REED, JANET E.  
 ; REGISTRATION NUMBER: 36,252  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (215) 563-4100  
 ; TELEFAX: (215) 563-4044  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3248 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: HUMAN  
 ; US-08-353-700-1

Query Match 3.58; Score 143; DB 1; Length 3248;  
 Best Local Similarity 19.5%; Pred. No. 0.017;  
 Matches 141; Conservative 110; Mismatches 265; Indels 208; Gaps 30;  
 QY 123 KEHGVIPRCODPRINELQOK-NLCTVVEVSYL-----EYIN-----E 162  
 DB 2159 KEKLLVRESQALSGESDYKLVNSKALEALVKEGFALRSLSTQEEVHQRGIE 2218  
 QY 163 RVROLNLSTGNLKV-----REHSPGVPVEDLAKLVRS--FQEIENLMDGKNA 212  
 DB 2219 KLRVTRDEKKLHIAEKLKERENDSLKDKVENRELQMSQENQELVILDAENKA 2278  
 QY 213 RTVAANTNMNNTSSSHAVFLITLQKWHDE--TKMDTEKVAKISLVDLAGSRATSTGA 270  
 DB 2279 EVELTKTOIEAMRSKLIFDLVLRSEKNTKQTEKQQLSELD-----KLSSFK 2333  
 QY 271 TGARLKEGAEI-----NRSLSTLGRVTAALADMSGKKKQLVPRUSVIT 317  
 DB 2334 SILLEKEQAEIQKEESTKAVEMIQNLKENENYALC-----GDQE----- 2376  
 QY 318 WLKDSLGSNGMTAMIAISPADINFEETLSTLRYADSARKIKHNHVVNEDNARMREL 377

DB 2377 -----IMKATQSILDP-PIEEHQI-----RNSIEKLRARLEADEKQKLCVLOQL 2420  
 QY 378 KEELAOLRSKLOSSGGGGGAGGSGPVEESYPPDTPLKQIVSLQOPDATVK--KMSKAE 436  
 DB 2421 KE--SEIHADLLK-----GRVE-----NLRELEIARTNQEHAALEANSKGE 2461  
 QY 437 I-----VEOLNQSEK-LVRDLNQTWEKLAKEPIHKEREAALEELGISIEKGFVGPVH 489  
 DB 2462 VETLKAKIBGTSQSLRGLDVTIRSEKENLUTNELQKEQ-RISLELEI-INSSPENILQ 2519  
 QY 490 SKEMPHLVNLSDDPELLACLIVYNIKPGOTRVGNVNDQAEIRLNGSKILKEHTF--EN 547  
 DB 2520 EKEQEKVQMKESSTAMEMLQTLKELNRYAALHNDQEA-----CKAKEQN 2566  
 QY 548 VDNVVTIVPEKKAAMVNGVRIDKPTRLSRGYRIILGDFHIFRHNPEEARAEQESLL 607  
 DB 2567 LSSQVECLEKKAQLQ--LDE--AKNYIVL-----QSSV 2599  
 QY 608 RHSVYNSQLSGPAPGRDHRTLSKAGSDGDSRSDSPLPHFGKSDWIFYARREASAIL 667  
 DB 2600 KGLIOEVEDGKOKLEKKDEEISRLKNQTDQEQVLVSKLSOVEGEHQLWKQONLELRNLTV 2659  
 QY 668 GLDQKISHL-----TDDELDALFDVQKARAVRRGLVEDNEDSDSGSPFYRDKYMNG 721  
 DB 2660 ELEQKIQVLQSKNASLQDTLEVL-----QSSY----- 2686  
 QY 722 TIDNFSLDITATMPGTPRSDDDGDALFFGDKSKQDASNVDFE---ELRQQAQMEAL 777  
 DB 2687 --KNLENELELTKM-----DKMSFEVKNVKMTAKETELQREMHMAQKTAELOEEL 2735  
 QY 778 KTAK 781  
 DB 2736 SGEK 2739  
 RESULT 12  
 PCT-US95-16216-1  
 ; Sequence 1, Application PC/TUS9516216  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yen, Timothy J.  
 ; APPLICANT: Rettner, Jerome B.  
 ; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently  
 ; TITLE OF INVENTION: Expressed Kinetochore Protein, and Methods of Use  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
 ; STREET: 1601 Market Street Suite 720  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19103-2307  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/16216  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/353,700  
 ; FILING DATE: 09-DEC-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Reed, Janet E.  
 ; REGISTRATION NUMBER: 36,252  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (215) 563-4100  
 ; TELEFAX: (215) 563-4044  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3248 amino acids

; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 PCT-US95-16216-1

Query Match 3.5%; Score 143; Db 5; Length 3248;  
 Best Local Similarity 19.5%; Pred. No. 0.017;  
 Matches 141; Conservative 110; Mismatches 265; Indels 208; Gaps 30;  
 QY 123 KHVGIPIRICODMFRINELQDK-NLTCVVSYL-----E1YN-----E 162  
 Db 2159 KEKELLVRESIQARLSDDYKLVSKALAEALVEKGFALRSSTOEVEHQLRGTE 2218  
 QY 163 RVLDLNPSTGKLV-----RHPSTGPGYVEDIAKLVR--FOEIMLMDENKA 212  
 Db 2219 KLVRIDEADEKQHLIAEKLERENDSLKQVLERELQWSENEQLVILDAENSKA 2278  
 QY 213 RYVAATNNETSSRSRAVFTLTQKHDEP--TKMDTEKVAKISLVLAGSERATSGA 270  
 Db 2279 EVETLKTQIEEMARSLKIFELDLVTLRSEKENLTKQIQEGQGLSELD-----KLSSFK 2333  
 QY 271 TGARLKEGAEI-----NRSLSLGRVIAALADMSSGKKKNOLVYPYRDSVLT 317  
 Db 2334 SLLEEKQAEIQIKESKTAVERMQLQKELNEVAALC-----GDQ----- 2376  
 QY 318 WLKDSLGNSKMTAMIAISPADINFEETLSLRYADSARKNHAVVNDPNARIMREL 377  
 Db 2377 LKVRIDEADEKQHLIAEKLERENDSLKQVLERELQWSENEQLVILDAENSKA 2278  
 QY 378 KEELAQLRSKLSQSGGGGAGGGGPGVPSYPPDTPLEKQIVSTQOPDATVK-KMSKAE 436  
 Db 2421 KE--SEHHAOLK-----GRVE-----NLERELIARTNQHAALAEANSKGE 2461  
 QY 437 I-----VEQLNQSEK-LYDOLNQTWEKLAKEETIHKERZAAELGISTEKGFGVGYH 489  
 Db 2462 VETLKAKEGMAQSLRLGLDVTYIRSEKENITNELQKEQ--RISELET--INSSFENILQ 2519  
 QY 490 SKEMPHLVNLSDDPLAECLVYNIKPGTAVGVYNDOTQAEIRLNGSKLKEHCTF--EN 547  
 Db 2520 EKEDEVOMKEKSTAMEMLOTOLKELNEVAALHDOEA-----CKAEQN 2566  
 QY 548 VQNVYTVIPNKAAYVNVGRIDKPTLRSGYRITLIDGHIFFRPHPEAARQESLL 607  
 Db 2567 LSSQVECLELXALQGL--LDE--AKNNIVL-----QSSV 2599  
 QY 608 RHVYTSNQLSPACRHDITLSKAGSDAGDSRSDSPHPFRGDSQWYFARAEASNL 667  
 Db 2600 KGLTQEVDCQKLEKKDEIRSLKNOIDQBLVSKLSQVGEHQLAKQCNLELRNLTV 2659  
 QY 668 GLDQKTSIL-----TDDELDFDQKARAVRGLVEDNEDSDSSQSFVRDQYNSNG 721  
 Db 2660 ELEOKTQVLOSKNLSQDTLEVL-----QSSV----- 2686  
 QY 722 TIDNFSLOTATMPCTPRDDGDALFFDKKSKODASNDVDE-----ELRQQAQMEAL 777  
 Db 2687 --KMLENELETKN-----DKMSFVKVNMKTAKETELQRMHEMAQKTAELQEEL 2735  
 QY 778 KTAK 781  
 Db 2736 SCEK 2739

RESULT 13

US-08-328-254-6  
 ; Sequence 6, Application US/08328254  
 ; Patent No. 5710022  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zhu, Xuelliang  
 ; APPLICANT: Lee, Wen-Hwa

; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell and Flores  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92122  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/328,254  
 ; FILING DATE: 24 OCT 1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/141,239  
 ; FILING DATE: 22-OCT-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Campbell, Cathryn A.  
 ; REGISTRATION NUMBER: 31,815  
 ; REFERENCE/DOCKET NUMBER: P-CJ 1191  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (619) 535-9001  
 ; TELEFAX: (619) 535-8949  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; LENGTH: 2482 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-328-254-6  
 Query Match 3.5%; Score 142; Db 1; Length 2482;  
 Best Local Similarity 19.6%; Pred. No. 0.014;  
 Matches 142; Conservative 108; Mismatches 266; Indels 208; Gaps 30;  
 QY 123 KHVGIPIRICODMFRINELQDK-NLTCVVSYL-----E1YN-----E 162  
 Db 1431 KEKELLVRESIQARLSDDYKLVSKALAEALVEKGFALRSSTOEVEHQLRGTE 1490  
 QY 163 RVLDLNPSTGKLV-----RHPSTGPGYVEDIAKLVR--FOEIMLMDENKA 212  
 Db 1491 KLVRIDEADEKQHLIAEKLERENDSLKQVLERELQWSENEQLVILDAENSKA 1550  
 QY 213 RYVAATNNETSSRSRAVFTLTQKHDEP--TKMDTEKVAKISLVLAGSERATSGA 270  
 Db 1551 EVETLKTQIEEMARSLKIFELDLVTLRSEKENLTKQIQEGQGLSELD-----KLSSFK 1605  
 QY 271 TGARLKEGAEI-----NRSLSLGRVIAALADMSSGKKKNOLVYPYRDSVLT 317  
 Db 1606 SLLEEKQAEIQIKESKTAVERMQLQKELNEVAALC-----GDQ----- 1648  
 QY 318 WLKDSLGNSKMTAMIAISPADINFEETLSLRYADSARKNHAVVNDPNARIMREL 377  
 Db 1649 -----TKATEQSLDP--PIEEHQL-----RNSIEKLARLEADEKQKLCVLOOL 1692  
 QY 378 KEELAQLRSKLSQSGGGGAGGGGPGVPSYPPDTPLEKQIVSTQOPDATVK-KMSKAE 436  
 Db 1693 KE--SEHHAOLK-----GRVE-----NLERELIARTNQHAALAEANSKGE 1733  
 QY 437 I-----VEQLNQSEK-LYDOLNQTWEKLAKEETIHKERZAAELGISTEKGFGVGYH 489  
 Db 1734 VETLKAKEGMAQSLRLGLDVTYIRSEKENITNELQKEQ--RISELET--INSSFENILQ 1791  
 QY 490 SKEMPHLVNLSDDPLAECLVYNIKPGTAVGVYNDOTQAEIRLNGSKLKEHCTF--EN 547  
 Db 1792 EKEDEVOMKEKSTAMEMLOTOLKELNEVAALHDOEA-----CKAEQN 1838

QY 548 VDNVTVTPNEKAAMVNGVRIDKPTRLRSGYRITLGLDHPHFRNHPPEARAERQBSLL 607  
DB 1839 LSSQVELELEKAQLLOG---LDE---AKNNYIVL-----QSSV 1871  
QY 608 RHSVTSNQLSGPAPGRDHTLRSKAGSDAGDSRSDSPHFRGKDSDFYARREASAIL 667  
DB 1872 NGLIQVEEDGQKLEKDEETSLRKNQIQDQBSLQVSGEGBHQLMKEQNLRLNIV 1931  
QY 668 GLDOKISHL-----TDELDALFDVQKARAVRGLVEDNEDSDOSSFFPVROKYMNG 721  
DB 1932 ELEQKIQVLQSKNASIQDTLEVL-----QSSY----- 1958  
QY 722 TIDNFSLDATITMPTGPRSDSDGALFEGDKSKODASNDVVE-----ELRQOQAMEAL 777  
DB 1959 --KNLENELELTRM-----DKMSFVKVKNMTAKETELOREHMAQKTAELQEEL 2007  
QY 778 KTA 781  
DB 2008 SKE 2011  
RESULT 14  
PCT-US93-03077-1  
; Sequence 1, Application PC/TUS9303077  
; GENERAL INFORMATION:  
; APPLICANT: Board of Regents, The University of Texas System  
; APPLICANT: Gaynor, Richard B.  
; APPLICANT: Wu, Foon Kin  
; TITLE OF INVENTION: PROTEIN CELLULAR FACTOR USEFUL FOR  
; TITLE OF INVENTION: REGULATING GENE EXPRESSION  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/03077  
; FILING DATE: 19930331  
; PRIORITY INFORMATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/07/862,025  
; FILING DATE: Apr 11, 2, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kammerer, Patricia A.  
; REGISTRATION NUMBER: 29,775  
; REFERENCE/DOCKET NUMBER: UFD270PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-787-1540  
; TELEFAX: 713-749-2679  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1093 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US93-03077-1

Query Match 3.5%; Score 140; DB 5; Length 1093;  
Best Local Similarity 18.1%; Pred. No. 0.0065;  
Matches 165; Conservative 119; Mismatches 325; Indels 302; Gaps 35;  
QY 1 MSGGNIKVVVRFPNAREIDRGAKCIVRNEG-----NQITLTPPGA--EEKARKS-- 51

DB 348 LSGKGYALVPIIVNGSTPK-----SKTVESAEGKSEVNETLVIPTPEAMESEGRSATP 402  
QY 52 -----GKTIIMD-----GPKAFAFORSVMSFOKNAPNARYQSD 83  
DB 403 VNCEQPDILVSTSPINEGQTVLDKVAEQCPAESQPEAL-----SEKED 446  
QY 84 LFQDLQV-----PLLDNAFGKYNCCIFAYGOTGSGKSYSMNGYKE 124  
DB 447 VCKTVEFLNEKLEKREAOALLSLSKKALLKEAFDLKDFMFRKEESS----- 495  
QY 125 HCVIPICQDMFRINELQDKMLCTCTVEVSLEYLYNERNVROLLNPSTGNLKVREHPT 184  
DB 496 -----ISSLKDEFTORIAEAEEKVOLAC-----KERDAARKEIK--NIKEELAT 537  
QY 185 GPVYEDLAKLVRSFOEINLMDENKARTVAATNNW-----ETSSSHAVFTLTITQW 239  
DB 538 RLNSSETADLLKEKDEOIRGLMEGEKLSKQOJHNSNIKKLRADKEMENNVAKLNKV 597  
QY 240 HDEETKMDTEKVAKISLVDLAGSERATSTGATCARLKEGAEINRSLSTLGRVIAALDMS 299  
DB 598 KELEELQHLQV-----LDGKEEVEKQHENIK--KLNSMVERQERDGLR--QVDM 647  
QY 300 SKQKKNQIVPYRDSV--LTWLLKDSLGSGNSMTAMTAATSPADINFEETLSLTLYADSA 356  
DB 648 ELEEKNRISQALDASAYKELTDLHKANAADS--EAPAAALSRMKAKKEELSALEKAOPE 706  
QY 357 KRKNHAVV-----NEDPNARMIRELKEELAOLRSKLSQGGGGGGAGGGGP 404  
DB 707 AROQVETLAIQVGDRLALQRTQAAARKEDYLHREIGELQORLOE-----AENRNOE 759  
QY 405 VEESYPPDT--PLEKQIVSIQDPDATVKKMSKAFIVEQLNQSEKLYRDL--NOTWEEKLAKT 462  
DB 760 LSQSVSTTRFLLRQENLQ-----ATLGSQTSWEKLEKLNLDRLGESOTLLA 808  
QY 463 EITHKREAALEBELGISIEKGFVGPYHS-----KEMPHLVNLSDDOPLLAEC 508  
DB 809 AAVERERATEELLANKTOMSSMESQNSILLRQNSRFQAOLESKNRCKLCKLENNRYQV 868  
QY 509 LVYNIKPGQTVGNVNDQTAQIEIRLNGSKILKEHCTFEN-----VDNVVTVIPNEKAAMV 563  
DB 869 ELEMLKDEYVR---TLEETRKETLLNSQLEMRKMKVEQERKKAIFTQETIEKERKPPS 925  
QY 564 VNGVRIDKPTPLKSGYRIILGDHPHFRNHPPEARAERQBSLLHHSVNSQLSGPAPGR 623  
DB 926 VSST-----PTMSKS-----SSISGVDMAGL 946  
QY 624 HDTLSKAGSDAGDSRSDSPHFRGKDSM--FYARREASATLGLD----- 671  
DB 947 OTSFLQ-----DESHDSFGMP---ISAKWKHJYACDKGSRKIKHENLQSLKLRG 998  
QY 672 KISHL-----TDELDALFDVQKARAVRGLVEDNEDSDSQ 708  
DB 999 EITHLOLEGNLEKFRSTMAEELYKLITNQDELEKVKYKPEIRLQJRLDLDQ----- 1050  
QY 709 SSPFPKVMKNGCTIDNFSLDATITMPTGPRSDSDGALFEGDKSKODASNDVVEELQ 768  
DB 1051 -----RY-----NTILQW-----YGERAEAEELRLDLELVKN 1078  
QY 769 O-QAOQMEALK 778  
DB 1079 WYKTQIDELLR 1089  
RESULT 15  
US-08-685-871-2  
; Sequence 2, Application US/08685871  
; Patent No. 6013499  
; GENERAL INFORMATION:  
; APPLICANT: NARUMIYA, Shuh  
; APPLICANT: IWAMATSU, Akihiro  
; TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160  
; NUMBER OF SEQUENCES: 68



---

Result	Query No.	Score	Query Length	ID	Description	
	1	1674	41.5	1150	1	kinesin-like prote
	2	1663.5	41.3	1695	1	kinesin family pro
	3	1593.5	39.5	1584	1	kinesin-like prote
	4	1590.5	39.5	1584	2	kinesin-like prote
	5	1396.5	34.7	1521	2	kinesin-like prote
	6	1384.5	31.9	1576	2	hypothetical prote
	7	1147	28.5	928	2	hypothetical prote
	8	1035	25.7	1121	2	kinesin-like prote
	9	1034	25.7	1121	2	kinesin-like prote
	10	860	21.3	699	1	kinesin-related br
	11	844.5	21.0	747	1	kinesin-related br
	12	828.5	20.6	742	1	kinesin-related br
	13	825.5	20.5	786	2	kinesin homolog
	14	814	20.2	955	2	kinesin-related pr
	15	800	19.9	701	1	LeKin kinesin-rela
	16	782.5	19.4	1231	2	kinesin-related pr
	17	775	19.2	1229	2	microtubule-associ
	18	772	16.2	928	2	kinesin-like prote
	19	765.5	19.0	793	2	kinesin heavy chai
	20	755.5	18.7	1325	2	kinesin-related br
	21	746.5	18.5	784	1	chromokinesin ch
	22	744	18.5	1226	2	kinesin-like prote
	23	743	18.4	935	2	kinesin (imported)
	24	722	17.9	929	2	kinesin [imported]
	25	708	17.6	672	2	kinesin osm-3 - Ca
	26	701	17.4	1130	2	kinesin heavy chai
	27	691.5	17.2	963	1	hypothetical prote
	28	689.5	17.1	294	2	kinesin-related br
	29	687.5	17.1	1031	1	kinesin heavy chai

QY	4	GMNIXVVRVAFNAREIDRGAQYIVBEGNOTILUTPPGAEKARKSGKTIWDGPKAPA	63
Db	1		11
QY	5	GASGVAVVRVFNRSRTSKESKCIITOMQGNSTISIPKNPKK-----APKSFS	51
Db	1		11
QY	64	FDRSYWSF--DKNAPNYARQEDIFQDLYGVLDDNAPKGYNANCIFYAGTQSGKSYMMGYG	122
Db	1		11
QY	52	FDYSYWSHTSPEDQCFASONRYINDIGKMLHAFQYINVCIFYAGTQSGKSYTMWGQK	111
Db	1		11
QY	123	KEHG--VPIRCDIMFRINELQDKNLCTIVESYLETYNRYEVRDLNAPNSTKGNKYRE	180
Db	1		11
QY	112	EESQAVIIFQCEELFEKIND--NCNEKMSYSEVSYMETYCYERVDLNPKNKGNLRVRE	170
Db	1		11
QY	181	HPSTGPFVDEKLALVYRSFQETENLMDGKNKARTVAATNMHSTSRSHAVFTLTLTQKH	240
Db	1		11
QY	271	HPILGPFVDEKSLAVYSTDLAOLMDAGNKARTVAATNMHSTSRSHAVFTVFTQKK	230
Db	1		11
QY	241	DEPTKQTEKVAKISLVDLAGSERASTGATGATKAGKEAGNEINSJLSTGRVYLAADNWS	300
Db	1		11
QY	231	DEPTWSTKSTSLVDLAGSERASTGATGRMEKAGNEINSJLSTGRVYLSLAESVK	290
Db	1		11
QY	301	GKQKNOLVPRDSVLTSLDKGNSNVAATSPADNPETLTSLTYADSKRQK	360
Db	1		11
QY	291	KKKTKDFIPYRDSVLTMLLENLGNCSNVAATSPADNPETLTSLTYADSKRQK	349
Db	1		11
QY	361	NHAYVVEDNEDNARIRELKPELQALRSKLGSSGGGGGAGCGGPVESTYPTPLEXQ--	418
Db	1		11
QY	350	CNAVIFEDNNAKILRELKEVTRIKQILRAQIGDITDPSNGSIT--SESSPSLSNQCV	407
Db	1		11

F: 6-354/Domain: kinesin motor domain homology <KMO>  
F: 97-104/Region: nucleotide-binding motif A (p-loop)

Db 350 CNAVINEDPNAKLVLRELKEEVTRLKDLLRAQGLGDIIDTSMGSLT--SSPSSCSLNSQVG 405

Db 169 REHPLGLPYVDLSKLAIVSYNDIQDMSGNKFRVVAATMMNTESSSHAVENIIITQK 228  
 QY 239 WHDETMMOTKVKAKISLVDLGASRATSGAGCARLKEAGENINSLSTLGRVIAALDW 298  
 Db 229 RHDATNITTEYKSKISLVDLGASRATSGAGTRKEAGANINSLTGLKVIKSLAEW 288  
 QY 259 SSG-----KQKNQLPVYRDSVLVWLLKDSLGNSWTAINAISPADINFEETLSLEYA 353  
 Db 289 DSGPNKKKKKTDIFIRDSVYLWLRNIGNSRTAKVAASPADINTDETSLURTA 348  
 QY 354 DSAKIKVHWAVNPDNAPWIKELKELAOLRSKLOSGGGGGAGSGGPPVEESYPPDT 413  
 Db 349 DRAKQIRCNAINEDPNKKLIPKLDVETRLDLYLAQGLG-----DIT 392  
 QY 414 PLEKOIVSIQODATVKKMS-----KATVEIQNSEKLYKROLNQW 455  
 Db 393 DMINALVGM-SPSSLSALSSRAASVSSLHERILFAPGSEERLERKETEKIHAELNETW 451  
 QY 456 EEKLAKTEETHKEREALBELGISLEK--GFWGYPHSEMPHLVNLSDPLAEKCLVYNI 513  
 Db 452 EEKLAKTEEAREREALLAEMGVAMREDGGTLGVFSPKKTTHLVNLNEDPLANECLLYI 511  
 QY 514 KPQOTRGVNYNDQTAERILNGSKILKEHCFEN-----VDNVYIVPEKAAVWNGVR 568  
 Db 512 KDSVTRVGREDAERQDVLVSGHIFKEEHCFIRSDSRGSGEAVVYLPESGADTVYNGKK 571  
 QY 569 IDKPTRLRSGYRIILGDFHIFRHPPEEARAFQBOQLSRHSVTSNQLSGPAPGRHRTL 628  
 Db 572 VTPLSLRNSGNRIINGKSHVFRHPHQARQER----- 606  
 QY 629 SKAGSDAGDSRSUSPLHFRGKDSOWFYARREASAILGLDQIKSHLTDDLDALFDV 688  
 Db 607 -----TPCAETPAEYVQWAFQARELLEK-QGDMK--QEMQRQLQLEDOY 649  
 QY 689 QKARAVRRGLVEDNDSQSFSFVRDKYMSN 720  
 Db 650 RRREBEATYLLR-QOURLYESKLEALQKOMS 680  
 RESULT 3  
 JN0114  
 Kinesin-related protein unc-104 - Caenorhabditis elegans  
 C:Species:Caenorhabditis elegans  
 C:Date:03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 29-May-1998  
 C:Accession:JN0114  
 R:Otsuka, A.J.; Jeyaparakash, A.; Garcia-Anoveros, J.; Tang, L.Z.; Fisk, G.; Hartshorne  
 Neuron 6, 113-122, 1991  
 A:Title:The C. elegans unc-104 gene encodes a putative kinesin heavy chain-like prot  
 A:Reference number:JN0114; PMID:91097805  
 A:Accession:JN0114  
 A:Molecule type:mRNA  
 A:Residues:1-1584 <CDS>  
 A:Cross-References:GB:M58582  
 A:Note:598-Thr and 930-Met were also found  
 C:Genetics:JN0114  
 A:Gene:unc-104  
 C:Gene family:kinesin-related protein unc-104; kinesin motor domain homology; pleckst  
 C:Superfamily:kinesin motor domain binding; e-loop  
 F:4-353/Domain:kinesin motor domain homology <KWOT>  
 F:91-100/Region:nucleotide binding motif (e-loop)  
 F:1285-1287/Region:cell attachment (R-G-n) motif  
 F:99/Binding site:ATP (Lys) #status predicted  
 Query Match 39.5%; Score 1593.5; DB 1; Length 1584;  
 Best Local Similarity 44.7%; Pred. No. 2,3e-78;  
 Matches 356; Conservative 139; Mismatches 213; Indels 89; Gaps 19;  
 QY 6 NIKVVRVFPNAREIDGAKIVRMGNOTILPPPPGAEBKARKSGKTMQGPAPAF 65  
 Db 3 SVKVAVRVFPNOREISNTSKCVLVNGNTTT-----NGHSINKENSFNED 50  
 QY 66 RSYWFSQKNAPNAPARQEDLFQDGLVPLLDNFAFKYNNCFAYGQTSKGSYSMMQYCK-- 123







C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16759  
R:Favell, T.  
submitted to the EMBL Data Library, March 1995  
A:Description: The sequence of C. elegans cosmid R144.  
A:Reference number: Z18571  
A:Accession: T16759  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-928 <FAY>  
A:Cross-references: EMBL:U23515; NID:g746492; PID:g746493; PIDN:AAC46545.1; CESP:R144.1  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:R144.1  
A:Introns: 80/3; 125/1; 261/2; 331/3; 350/2; 516/3; 589/1; 615/2; 638/3; 726/1; 865/3; 928/1  
Query Match 28.5%; Score 1147; DB 2; Length 928;  
Best Local Similarity 38.8%; Pred. No. 2, 4e-54;  
R:Map position: 2  
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology  
Matches 281; Conservative 116; Mismatches 217; Indels 110; Gaps 19;  
Qy 1 MSGGGINVYVVRFPNAREIDRGAKCIVRMGGNOTILTPPGAEKARKSGKTMDCPK 60  
Db 1 MKGDSIIIVAVRFPNDREXTRNCKLVTEMDETTVIRPKTNDE-----K 48  
Qy 61 AFAPDRSYWFD-----KNA-----PNYARQEDFDGLGVPLLDNAFGYNNCIFAYGQ 109  
Db 1 AFAPDRSYWFD-----KNA-----PNYARQEDFDGLGVPLLDNAFGYNNCIFAYGQ 109  
Qy 49 RTVDHSYWSHGDTSEKNGVLEPTDHPADORRVEDFGLRGVLANAWAGVNSLFPAYGQ 108  
Db 49 RTVDHSYWSHGDTSEKNGVLEPTDHPADORRVEDFGLRGVLANAWAGVNSLFPAYGQ 108  
Qy 110 TGSCKSYSMNGYKGHEGVIPRCQMPFRINELQKOKNLCTVEYSLEINERVLDLN 169  
Db 110 TGSCKSYSMNGYKGHEGVIPRCQMPFRINELQKOKNLCTVEYSLEINERVLDLN 169  
Qy 109 TGSCKSYSMNGYKGHEGVIPRCQMPFRINELQKOKNLCTVEYSLEINERVLDLN 167  
Db 109 TGSCKSYSMNGYKGHEGVIPRCQMPFRINELQKOKNLCTVEYSLEINERVLDLN 167  
Qy 170 --PSTGNLKVREHSPSTGYDLAKLVRSPOEINLADGNGKARTVAATNWFSSRS 227  
Db 170 --PSTGNLKVREHSPSTGYDLAKLVRSPOEINLADGNGKARTVAATNWFSSRS 227  
Qy 168 STTPPKGLAVRHPKNGFYENLITVYVNSKELAEIGTSRTIAATQWATSSRA 227  
Db 168 STTPPKGLAVRHPKNGFYENLITVYVNSKELAEIGTSRTIAATQWATSSRA 227  
Qy 228 HAVFTLTITQKHDEETKMDTERVAKLSIVDLGSRATSTGATGARKLGAERISLT 287  
Db 228 HAVFTLTITQKHDEETKMDTERVAKLSIVDLGSRATSTGATGARKLGAERISLT 287  
Qy 228 HTVKTITPNOK--SSQAGGTSMMKSEINLVLAGSRQSAAGTECDRLKEGIVNOSLTT 286  
Db 228 HTVKTITPNOK--SSQAGGTSMMKSEINLVLAGSRQSAAGTECDRLKEGIVNOSLTT 286  
Qy 288 LGRVIAALADKSGKQKQNLVPRYDSVTLWLLKSLGNSMTAMIAISPADINFEETL 347  
Db 288 LGRVIAALADKSGKQKQNLVPRYDSVTLWLLKSLGNSMTAMIAISPADINFEETL 347  
Qy 287 LGRVIALHDSQAKSCKRTQIPYRDSVTLCLKNALGNSKTMIAISPADINFEETL 346  
Db 287 LGRVIALHDSQAKSCKRTQIPYRDSVTLCLKNALGNSKTMIAISPADINFEETL 346  
Qy 348 STLYADAKRKTKHAYVNEPNAIRTELKPELAQLRSKLGSGGGGGAGSGGPVEE 407  
Db 348 STLYADAKRKTKHAYVNEPNAIRTELKPELAQLRSKLGSGGGGGAGSGGPVEE 407  
Qy 347 STLRADAKRKTKHAYVNEPNAIRTELKPELAQLRSKLGSGGGGGAGSGGPVEE 393  
Db 347 STLRADAKRKTKHAYVNEPNAIRTELKPELAQLRSKLGSGGGGGAGSGGPVEE 393  
Qy 408 SYPPDPTFLEQIVSTOORDATVKKSKAEIVDLQNGSEKLYKSLNQTWEEKLAKTEEIK 467  
Db 408 SYPPDPTFLEQIVSTOORDATVKKSKAEIVDLQNGSEKLYKSLNQTWEEKLAKTEEIK 467  
Qy 394 -----DASNEIEK--LRQLAENKQEWKESKQKRIA--EENAK 431  
Db 394 -----DASNEIEK--LRQLAENKQEWKESKQKRIA--EENAK 431  
Qy 468 EREAALLEGISIEKFGVPGVBSKEMPHLVNSDDFLAECVYNKPKGQTEGVNVD 527  
Db 468 EREAALLEGISIEKFGVPGVBSKEMPHLVNSDDFLAECVYNKPKGQTEGVNVD 527  
Qy 432 HASGASEKVEAMEAKK-----KMKCHLNINDPALTIVVHFIPVGESVGNKPTSS 483  
Db 432 HASGASEKVEAMEAKK-----KMKCHLNINDPALTIVVHFIPVGESVGNKPTSS 483  
Qy 528 QAEIRLNSKILKEHCTFEN--VONVTVTP--NEKAAVMVNGYRIDKPTFLASGYRIIGD 585  
Db 528 QAEIRLNSKILKEHCTFEN--VONVTVTP--NEKAAVMVNGYRIDKPTFLASGYRIIGD 585  
Qy 484 GNFTONSGSLIPQVTLKNDGNNGNIIHLSPCEDLDIFINGKPHIGETLOAQNDVFFVG 543  
Db 484 GNFTONSGSLIPQVTLKNDGNNGNIIHLSPCEDLDIFINGKPHIGETLOAQNDVFFVG 543  
Qy 586 PFIFRNHPEARAEOROSILRHGVTNSQLGSPAPGRDRTLKAGSDAGDGRSD--- 642  
Db 586 PFIFRNHPEARAEOROSILRHGVTNSQLGSPAPGRDRTLKAGSDAGDGRSD--- 642  
Qy 544 NHLYVFNPNKTKG-----IHTDITENAAQAEIAQHAAALNGR--LGGSKRDLIL 593  
Db 544 NHLYVFNPNKTKG-----IHTDITENAAQAEIAQHAAALNGR--LGGSKRDLIL 593  
Qy 643 -----SPLPFHGKDSWFYAREASAIILGDKLIS--HUTDDELALFDVQKARVR 696  
Db 643 -----SPLPFHGKDSWFYAREASAIILGDKLIS--HUTDDELALFDVQKARVR 696  
Qy 594 EEELANSTLP-----LVORANNAIYELGRNKEIVLVSPEN-----RGITS 634  
Db 594 EEELANSTLP-----LVORANNAIYELGRNKEIVLVSPEN-----RGITS 634  
Qy 697 GLVE 700  
Db 697 GLVE 700  
Qy 635 GLTE 638  
Db 635 GLTE 638

RESULT 8  
T13750  
linesin-like protein 388 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C:Accession: T13750  
R:Ohkura, H.; Torok, T.; Tick, G.; Bohetsel, J.; Kiss, I.; Glover, D.M.  
J. Cell Sci. 110, 945-954, 1997  
A:Title: Mutation of a gene for a Drosophila kinesin like protein, KLP388, leads to f  
A:Reference number: Z17752; MUID:97296456  
A:Accession: T13750  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-1121 <OHK>  
A:Cross-references: EMBL:Y10667; NID:g253151; PIDN:CAAT71675.1; PID:g2253152  
C:Genetics:  
A:Cross-references: Flybase:FBgn0004374  
A:Map position: 2  
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology  
Matches 275; Conservative 131; Mismatches 293; Indels 116; Gaps 21;  
Query Match 25.7%; Score 1035; DB 2; Length 1121;  
Best Local Similarity 33.7%; Pred. No. 2, 4e-48;  
R:Map position: 2  
C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology  
Matches 275; Conservative 131; Mismatches 293; Indels 116; Gaps 21;  
Qy 1 MSGGCKTKVYVVRFPNAREIDRGAKCIVRMGGNOTILTPPGAEKARKSGKTMDCPK 59  
Db 115 VSEESNIVAVRVPRLNALLECTRGQVTNVQVHGNSNELTVQAGSSADA--SAGVT----- 168  
Qy 60 KAPAFDRSYWFDKNAIPYARQEDFDGLGVPLLDNAFGYNNCIFAYGQTSKGSYNNM 119  
Db 60 KAPAFDRSYWFDKNAIPYARQEDFDGLGVPLLDNAFGYNNCIFAYGQTSKGSYNNM 119  
Qy 169 HFFSYDGVYVSCOPKRNKFAQCVFEGTARPLIDTAPRGYNACLPAFGQTSKGSYNNM 228  
Db 169 HFFSYDGVYVSCOPKRNKFAQCVFEGTARPLIDTAPRGYNACLPAFGQTSKGSYNNM 228  
Qy 120 GY-----GKEH--GVTPICQDFRINELQKOKNLCTVEYSLEINERVLD 166  
Db 120 GY-----GKEH--GVTPICQDFRINELQKOKNLCTVEYSLEINERVLD 166  
Qy 229 GTEALDDAALGGPDRHEAGIIPRFCHLEFRLEAVKSOQLQVEVYSLEINERVLD 288  
Db 229 GTEALDDAALGGPDRHEAGIIPRFCHLEFRLEAVKSOQLQVEVYSLEINERVLD 288  
Qy 167 LLN-----PSTGNLKVREHSPSTGYDLAKLVRSPOEINLADGNGKARTVAATNWFSSRS 208  
Db 167 LLN-----PSTGNLKVREHSPSTGYDLAKLVRSPOEINLADGNGKARTVAATNWFSSRS 208  
Qy 289 LLSVQHAAATGESTPIQ000000PALKVREHPIFGVYVYDLISAHSYDSYALRNKLAV 346  
Db 289 LLSVQHAAATGESTPIQ000000PALKVREHPIFGVYVYDLISAHSYDSYALRNKLAV 346  
Qy 209 GNKARTVAATNMETSSRSYHVF--TLTITQKHDEETKMDTERVAKLSIVDLGSRATSTGATGARKLGAERISLT 256  
Db 209 GNKARTVAATNMETSSRSYHVF--TLTITQKHDEETKMDTERVAKLSIVDLGSRATSTGATGARKLGAERISLT 256  
Qy 349 GNSQRTASTAMNDKSSRESHFVNLNLTDLSSDGLSDTSDTSASSLRQTRRSKLSL 408  
Db 349 GNSQRTASTAMNDKSSRESHFVNLNLTDLSSDGLSDTSDTSASSLRQTRRSKLSL 408  
Qy 257 VDLGASERATSTGATGARKLGAERISLTSLGRVIAALADKSGS-----GKQKNQV 309  
Db 257 VDLGASERATSTGATGARKLGAERISLTSLGRVIAALADKSGS-----GKQKNQV 309  
Qy 409 VDLGASERATSTGATGARKLGAERISLTSLGRVIAALADKSGSNGPGLSGTSTFV 468  
Db 409 VDLGASERATSTGATGARKLGAERISLTSLGRVIAALADKSGSNGPGLSGTSTFV 468  
Qy 310 PRDSVTLWLLKSLGNSMTAMIAISPADINFEETLSTLYADAKRKTKHAYVNEP 369  
Db 310 PRDSVTLWLLKSLGNSMTAMIAISPADINFEETLSTLYADAKRKTKHAYVNEP 369  
Qy 469 PTRESVTLWLLKSLGNSMTAMIAISPADINFEETLSTLYADAKRKTKHAYVNEP 528  
Db 469 PTRESVTLWLLKSLGNSMTAMIAISPADINFEETLSTLYADAKRKTKHAYVNEP 528  
Qy 370 NARMTRELEAQLRSKLGSGGGGGAGSGGPVEEYPTPLEKQIVSTOORDATV 429  
Db 370 NARMTRELEAQLRSKLGSGGGGGAGSGGPVEEYPTPLEKQIVSTOORDATV 429  
Qy 529 HDKIIIRDLRAEVDRKLSRNEYERQBRSLGNSNPNVPRKIET-----SVDETEVEA 581  
Db 529 HDKIIIRDLRAEVDRKLSRNEYERQBRSLGNSNPNVPRKIET-----SVDETEVEA 581  
Qy 430 KMKSKAEIVDLQNGSEKLYKSLNQTWEEKLAKTEEIKREAALELGLSIEKFGVPGV 489  
Db 430 KMKSKAEIVDLQNGSEKLYKSLNQTWEEKLAKTEEIKREAALELGLSIEKFGVPGV 489  
Qy 582 LRQQLAERERELSAQK-----SMWEKLEAEDORKSELVLRKRLALEL-----TA 629  
Db 582 LRQQLAERERELSAQK-----SMWEKLEAEDORKSELVLRKRLALEL-----TA 629  
Qy 490 SKEMPHLVNSDDFLAECVYNKPKGQTEGVNVD-----NYNQDTQAEIRLNSKILKEHCT 544  
Db 490 SKEMPHLVNSDDFLAECVYNKPKGQTEGVNVD-----NYNQDTQAEIRLNSKILKEHCT 544  
Qy 630 EQKQNCVLWLTADPILSGTLFYLPGQVLRIGRGLPGGSSSQDPIVDGLGFLVAQLHCS 689  
Db 630 EQKQNCVLWLTADPILSGTLFYLPGQVLRIGRGLPGGSSSQDPIVDGLGFLVAQLHCS 689  
Qy 545 FENV--ONVTVTPNEKAAVMVNGYRIDKPTFLASGYRIIGDPIFFHNP-----EENAR 599  
Db 545 FENV--ONVTVTPNEKAAVMVNGYRIDKPTFLASGYRIIGDPIFFHNP-----EENAR 599  
Qy 690 IEREGKQTLFGPESDFTYVNGYELKQRQLFHWGDLRVIGGSHFYRISNPFCSQRGA 749  
Db 690 IEREGKQTLFGPESDFTYVNGYELKQRQLFHWGDLRVIGGSHFYRISNPFCSQRGA 749  
Qy 600 EQKQNCVLWLTADPILSGTLFYLPGQVLRIGRGLPGGSSSQDPIVDGLGFLVAQLHCS 649  
Db 600 EQKQNCVLWLTADPILSGTLFYLPGQVLRIGRGLPGGSSSQDPIVDGLGFLVAQLHCS 649

Db 750 DHPVDFQLAHOELLQKQQLRSELEAEKRAALTKIEQERAOHARDFEERLQCLELQF 809  
 QY 650 GK-DSDFYARREASAILGDKDISHLTDELDALEDDVOKA---RAVRRLGVEDNEDS 705  
 Db 810 YKCNSEMLETERQA---LALAQQOEHPLRHEDAVSTPAQKSTILEDIORIMLPNSES 865  
 QY 706 DSQSSFPVRD-----KYSNGTIDNFSL 728  
 Db 866 LHKTLMVKEATQRCQLDPLFEFRQTQDPDFGL 900  
 RESULT 9  
 T13796  
 Kinesin-related protein 38B - fruit fly (*Drosophila melanogaster*) (fragment)  
 C:Species: *Drosophila melanogaster*  
 C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C:Accession: T13796  
 R: Molina, I.; Baars, S.; Hales, K.; Fuller, M.T.; Ripoll, P.  
 J. Cell Biol. 139, 1361-1374, 1997  
 A:Title: A chromatin associated kinesin-related protein required for normal mitotic chromosome segregation  
 A:Reference number: z17759; MUID:98060833  
 A:Accession: T13796  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1121 <MOL>  
 A:Cross-references: EMBL:Y15247; NID:g2578010; PIDN:CAA75531.1; PID:g2578011  
 C:Genetics:  
 A:Gene: tio  
 A:Cross-references: FlyBase:FBgn0004374  
 C:Superfamily: kinesin-related protein E95; kinesin motor domain homology

Query Match 25.74; Score 1034; DB 2; Length 1121;  
 Best Local Similarity 33.8%; Pred. No. 2.7e-48;  
 Matches 274; Conservative 130; Mismatches 290; Indels 116; Gaps 21.

QY 6 NIKYVVRFPFNAEIDRG-AKCIVRMEGNOTILTPPGAEEKARKSGKIMDGPKAFAP 64  
 Db 120 NMIVAVVRNPNALECTGGTQVTVVGVHNSLTVQAGSSADA-SAGVT-----HEFSY 173  
 QY 65 DRSWFSFKNAPNVARQEDLPQDLGVPLLDNAFKYNCIFAYGQTSGKSYSMGY---121  
 Db 174 DQVYISCDPERKFAQAKVEGTARPIDYAFEGYNACLAFAYGQTSGKSYSMGIEAL 233  
 QY 122 -----GKEH---GVPIRCQDMFRINELQDKNLTCTVEVSLIYNERVRDLLN--169  
 Db 234 DDAALGGPPHYENGIFPFCHELFERLEAVKSOQLQVEVSYFELYNKIHLLSVQ 293  
 QY 170 -----PSTGKNLKVREHPSTGPIYEDLAKLVRSFQETENLMDGKNKAR 213  
 Db 294 HAAATGESTPIQQQQOQORQALKVREHPITGPIYEDLSAHSYDSYSAURLNVLAVNSOR 353  
 QY 214 TVATNNETSSSHAVF--TLTLTKQKHDETKMDTE-----KVAKISLDVLQ 261  
 Db 354 ATASTAKNDSSKSHSINIVLMLDLSDDGLSSDDYDSTASSIRTRKSKISLDVLQ 413  
 QY 262 SERATSGTGAELKGAENRSLTGLGVIAALADMS-----GKQKNQLVPPROS 314  
 Db 414 SERLSVSGNERIGVSNKSLTGLKVIKALADSKAANGPLGSGTPTVPYRES 473  
 QY 315 VLTWLLKSLGNSMTAMIAISPADINFEETLSLTVADSAKRIKHAIVNEDNAPMI 374  
 Db 474 VLTWLLRNLGNSKTVMLATISPASIHDETALITLACKASIVRVKYNVESPHDKII 533  
 QY 375 RELKEELAKLSQSGGGGCGGAGGVSEYPPDPLEKQIVSQIDQDPAIVKMSK 434  
 Db 534 RDLRAEVDRKLRLNEVQRRLSGNSNPNVPKIILET-----SVDETEVALRQQL 586  
 QY 435 AEIVQOLNOSKLYRDLNQTWEKLAETKEHREKAALEELGISIEKFGVGYHSEMP 494  
 Db 587 AERERELSAQK-----SMWEKLEAKEDORSRLYLRKRGALAEI-----TAEDKQA 634  
 QY 495 HLNVLSDDPLLAELCLVYNIKPGQTRVG-----NVNQDTQAEIRLNGSKILKEHCTFENV- 548

Db 635 CLIVNLTAQDPIILSGTLYLLPGLVRIIGRGLPGSSSSQPDIVLDGFLVALQCHSIEHR 694  
 QY 549 -DNVVTIVPNEKAAMVWNCVRIKDPTRLSRGYRIILGDPIHFRNHP---BEAARAEQEO 604  
 Db 695 GCKLVIPGSGEPETVYNGELLKDRQLFHGDRLVIGGSHYFRISNPFCSQRSGADHPVD 754  
 QY 605 SILRH---SVTNSQSGPAPGRHRTLSKAGSDADGDSRDS-----PLPFRGK-DS 653  
 Db 755 POLAHOELLQKQQLRSELEAEKRAALTKIEQERAOHARDFEERLQCLELQFQKYKNS 814  
 QY 654 DMFYARREASAILGDKDISHLTDELDALEDDVOKA---RAVRRLGVEDNEDSOS 710  
 Db 815 EMLETERQA---LALAQQOEHPLRHEDAVSTPAQKSTILEDIORIMLPNSESLLHKTQ 870  
 QY 711 FVRD-----KYSNGTIDNFSL 728  
 Db 871 LMVKEATQRCQLDPLFEFRQTQDPDFGL 900  
 RESULT 10  
 S38982  
 Kinesin-related protein KRP85 - sea urchin (*Strongylocentrotus purpuratus*)  
 N:Alternate names: kinesin-2 chain A; KRP (85/95) 85k chain  
 C:Species: *Strongylocentrotus purpuratus* (purple urchin)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: S38982; S72551  
 R: Cole, D.G.; Chinn, S.W.; Wedaman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.  
 Nature 366, 268-270, 1993  
 A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.  
 A:Reference number: S38982; MUID:94050179  
 A:Accession: S38982  
 A:Molecule type: mRNA  
 A:Residues: 1-699 <COLI>  
 A:Cross-references: EMBL:L16993; NID:g295245; PIDN:AAA16098.1; PID:g295246  
 A:Accession: S72551  
 A:Molecule type: protein  
 A:Residues: 2-5,'X',7-11;59-64;125-132;222-226,'X',228-230 <COL2>  
 C:Complex: heterotrimer of a 115k chain and two kinesin-related chains of 95k (PIR:S5)  
 C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
 C:Keywords: ATP; heterotrimer; microtubule binding; P-loop  
 F:11-348/Domain: kinesin motor domain homology <KMOR>  
 F:97-104/Region: nucleotide-binding motif A (P-loop)  
 F:103/Binding site: ATP (Lys) #status Predicted

Query Match 21.34; Score 860; DB 1; Length 699;  
 Best Local Similarity 40.04; Pred. No. 3.7e-39;  
 Matches 205; Conservative 88; Mismatches 155; Indels 64; Gaps 15.

QY 2 SGGGNKVVVRFPFNAEIDRGAKCIVRM-EGNOTILTPPGAEEKARKSGKIMDGPK 60  
 Db 6 SGNQNVVVRVRCPLNKSKEQTGGFKSVVKMDKRGTVQVTNPNA-----PSG---EPPR 56  
 QY 61 AFAPDSYWSFDKMANVARQEDLPQDLGVPLLDNAFKYNCIFAYGQTSGKSYSMG 120  
 Db 57 SFEDTVF-----AEG-AKTQDVYNOTARPVIDALIEGVTGFIAYGQTGKTFTFMG 109  
 QY 121 YGKE---HGVPIRCQDMFRINELQDKNLTCTVEVSLIYNERVRDLLNPSTKGNK 177  
 Db 110 VNSPELRLGILPNSFAHIFGH--AKEQNVRLVRSYLEIYNEVKDLKCKQOQHRL 167  
 QY 178 VREHPSTGPIYEDLAKLVRSFQETENLMDGKNKARTVATNNETSSSHAVFTLQ 237  
 Db 168 VKEPQGVYVVDLSAFVNNADDMRTLTKNSRVGATNNSSRSRHAFTFLER 227  
 QY 238 KWHDETKMDTE---KVAKISLDVLQSGERATSGTGAELKGAENRSLTGLGVIAA 294  
 Db 228 -----SDWLKDEQHVQRKLIHWLDLQSGERQTKGTGATGQRLKATKINLSLTGLNVISS 283  
 QY 295 LADMSGKQKNQOLVPYRDSVLTWLLKSLGNSMTAMIAISPADINFEETLSLTVAD 354  
 Db 284 LVD-----GKSTHPIYRNSKLTLLQSLGNGNAKTVMCANIGPAEYNTDETISTLAYN 337

336 RAKNTKRPVNEDEPKDALURFQETELAKQAOLBKRSIGRRKRKRREGGSGGGGEE 395

405 VE---ESYPPTPLEQTIVSQDPATKYKSRRAEIVP--QLNOSKEKIVRLDNLWPEKL 459  
          ::: :         ::: :         ::: :         ::: :  
396 EBECEGECEGDODKDIDYWRQEKELEIKRA---IYEDHSILHVLNLSDDPLLAECLY 511  
          ::: :         ::: :         ::: :         ::: :  
460 AKTEIHKKERAAELGLGISIE---KGPGVPY---HSEKMPHLVNLSDDPLLAECLY 511  
          ::: :         ::: :         ::: :         ::: :  
448 KKHEDLRREKDAE-EMLCAMIKAMESLLGVGNKVDTHTNEQQILEKQRQETARQ--- 502  
          ::: :         ::: :         ::: :         ::: :  
512 NIKPGQTRVGVNQDQTAELIRNSKLIKHECTFENVDNVVTIPNKAAYVINGVRDK 571  
          ::: :         ::: :         ::: :         ::: :  
503 -----KRREREIQO-QMESRDETLEIKE-TVTSLQQEQDI-----KTK 540  
          ::: :         ::: :         ::: :         ::: :  
572 PTRLSRGYRIILDHPHFRENHPEAERBQS-----LLRHVSINSOLG----- 617  
          ::: :         ::: :         ::: :         ::: :  
541 LKLPFSKLAQVAEIIHQDEEHIREQELETONEIFRELKULKIITFIPLKEKNIM 600  
          ::: :         ::: :         ::: :         ::: :  
618 -----SPAGRHURTSLKAGSDADGSDS-----PLPIHERKDS 653  
          ::: :         ::: :         ::: :         ::: :  
601 NRSPFDDEDHMKLPHITLENQNMKRPVSNGVKRPLSQLAHMWSMIIPERYAEH- 659  
          ::: :         ::: :         ::: :         ::: :  
654 DMFYAREAAALIGLKISHLTDELDAFLDDVKARAHRGLGVNFDSQSSSF 711  
          ::: :         ::: :         ::: :         ::: :  
660 -----IMLLDELPSRTTRDYEGPAISPKVO--AALDAALOQDEIDEVDASSF 705  
          ::: :         ::: :         ::: :         ::: :  
  
RESULT 12  
S58691  
kinesin-related protein KRP95 - sea urchin (Strongylocentrotus droebachiensis)  
N:AItername names: kinesin-2 chain B, KRP (85/95) 95K chain  
C:Species: Strongylocentrotus droebachiensis  
C:Cdate: 10-Sep-1999 fsequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
Accession: S58691  
RefSeq: NM\_015462.1 Genbank: F045666.1 PIR: P45666  
J Mol Biol. 251(2):157-162, 1995  
A>Title: Heterodimerization of the two motor subunits of the heterotrimeric kinesin.  
A:Reference number: S58691; MUID:95404610  
A:Accession: S58691  
A:A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-742 <RAS>  
C:C:Complex: heterotrimer of a 115K chain and two kinesin-related chains of 85K (PIR:P45666)  
C:C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
C:Keywords: ATP; heterotrimer; microtubule binding; P-loop  
F:9-345/Domain: kinesin motor domain homology <KMOT>  
F:F95-102/Region: nucleotide-binding motif A (P-loop)  
F:F101/Binding site: ATP (lys) #status Predicted  
  
Query Match 20.6%; Score 826.5; Db 1; Length 742;  
Best Local Similarity 31.5%; Pred No. 26-37;  
Matches 250; Conservative 122; Mismatches 286; Indels 135; Gaps 25;  
  
QY 7 IKVVYRVRFNAERDGAKCIIVREMGNTI--LTPPGAEEKARKSKCTIMGDPKAFAP 64  
DB :::::::::::::::::::::  
QY 9 KVYVVYRVCRPMASKEISQHKRIIVEMONKRGIVNPVFKGPCEPNKS-----FTF 58  
DB :::::::::::::::::::::  
  
QY 65 DRSY-SFDFKANPVARQEDFDQLGVPLLDNAFNKYNICIFYGTGSGKSYSMNCYK 123  
DB :::::::::::::::::::::  
QY 59 DTAVYDWN-----SKOIDYDFTSRISVESVLQGTFNIFYGYTGTCKITFMCGVRS 110  
DB :::::::::::::::::::::  
  
QY 124 E---HGVIPICDMFRIINELQDKDMCTIVEVSLYEINERVRLDLNLNSTKNGLVKRE 180  
DB :::::::::::::::::::::  
QY 181 HPSTGPGVEDLAKLVAFSEOTEHLADECNKARTVANTWMTFSRSSNIATFLTLQOKHW 240  
DB :::::::::::::::::::::  
QY 166 RDGTGVYKVALSSFTYSKVELERTVTVGNNSRVTGNNSSRSHSIPIITI----E 237  
DB :::::::::::::::::::::  
  
QY 241 DEETKMDTE--KVAKLSLVDLAGSERATSTGATGARKEGATEINPSLSTGRIVALAD 293  
DB :::::::::::::::::::::  
QY 224 CSELGVGDGNIHVWKLNVLVDLAGSERQAKTGATGDRLEAKTKINLSLALGNYSALVD 283  
DB :::::::::::::::::::::

```

QY 298 MSSGOKKQNLVPRDYSVLTKLSDSGNSMTAMIAISPADINFEETSLSTRYDSAK 357
DB 284 -----GSSSHIFPRDLSKLRLLQSDGGWAKTVANMCPASYNPDITITLYANRAK 337
QY 358 RIKNHAVNEDPNARIMRELKELAEQRLSKLQSSG-GGGGAGGAGGPGVEESYPPOTPLE 416
DB 338 NIKNPKPNEDPDALLREFQESIRLQKALQKGPQSDGRKKRKGKPPGEQGGDDIEDE 397
QY 417 KQIVSQIQPDATVKKMSKAEIVFQNLNSSEKLYARDJNOTWEEKLAKTET-----H 466
DB 398 TEEEGDEMEDEMYKESQQLKEE--EKELMANQSMIAEAEKQALLSEVOKRQGEIKKEH 454
QY 467 KEREAAELGIGIEKGPGVGY-----HSKEMPLVNLSDDPDLAECLVYNIKPGQTRVGN 522
DB 445 QOKEMLEGIKAMESKLLVGGKSIYVDHTNQORKEE-EQRLLLAEE-----EKN 501
QY 523 VNQDQQAIRLNGSKILKEICHTFENVQNVVTVIPNKEAAVMVNGVRIDKPTRLRSYRII 582
DB 502 REROMERLKBODDKTVTEIGTFSSLOQVE-----VTKTKLKLKFLAQSY 548
QY 583 LGDFHIFRNPHEPAR-AERQESQLSRHSVTSQSGPAPGRHDT--LSRAGSDADGDS 639
DB 549 KSDIQDQDQEHAREQELQTONELIRELKLKVVTAONFTPVPEENTKITTRAVFDETEE 608
QY 640 RSDSPULPIHFGKD-----SDWFYAR-----REAAAILGLQKIS 674
DB 609 WLTLPLAKAGSPQMAKRPVSAGVGNRRPIAD--YARMAQMGPNRYKAEINILSVLDMP 666
QY 675 HLT--DDELDALFDQVKARAVRGLVEDNEDSDQSSEFV-----RKYMSNGTID 724
DB 667 NMTTRDYGSPVAPRQV---AALDAALQDEDDLDLEQVEPFKAKTKLKKDKVRSK---- 719
QY 725 NFSLDTAITMPT 737
DB 720 -----HRAVAKPGS 728

RESULT 13
A:Species: Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 10-Dec-1999
C:Accession: A53939
R:Walther, Z.; Vashishtha, M.; Hall, J.L.
J. Cell Biol. 126, 175-188, 1994
A:Title: The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous protein.
A:Reference number: A53939; MUID:94299638
A:Accession: A53939
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-786 <BL>
A:Cross-references: BMEL:I33697; NID:9497696; PIDN:AAA21738.1; PID:9497697
A:Note: Authors translated the codon AAC for residue 753 as Asp
C:Genet: FLA10
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; coiled coil; P-loop; homology <KMOT>
F:11-359/Domain: kinesin motor domain homology <KMOT>
F:97-104/Region: nucleotide-binding motif A (P-loop)

Query Match 20.5%; Score 825.5; DB 2; Length 786;
Best Local Similarity 39.3%; Pred. No. 3, 2e-37;
Matches 204; Conservative 80; Mismatches 168; Indels 67; Gaps 14;

QY 2 SGGS--NIKVVVVRFPNAREIDRCACIVRME---GNQITLTPPGAEEKARKSGKTI 56
DB 4 AGGSGSEKVVVRCPRLNGEKADGRSRIYDMVDQAVKVRNPKADASEP----- 54

QY 57 DGPKAFAPDRSY-WFSDKNAPNARYQEDLPDGLVPLLDNAFKYNNCIFAYGQTSGKS 115
DB 55 --KRAFTDQVDMN-----CQQRDVFDTARPLDSCIEGYNTIFAYGQTSGKS 104

```

```

QY 116 YSMNGYKRE--HGVIPRIQDMFRKINELQKKNLCTVEVSYLEIYNERVROLLNPST 172
DB 105 HTMEKDEPPPLGLPNTFRYVFEIATDSGTKEF--LVRSSYLEIYNERVROLLGKH 162
QY 173 KGNLKVHPHSTPGYVEDLAKLVRSFOETENLMDGKNKARTVAATNMNETSSRSHAVFT 232
DB 163 SKKMKELKSPDRGVVYKQLSQFYCKNVEMNKKVLLAGKDNQVQVATLMNQDSSHSIET 222
QY 233 LTL-----TQKWHDBETKMTKEVAKISLVDLAGSERATSTGATGARLKEGAETN 282
DB 223 ITTECTEKLESAQAOKPGAKKDDSHVRVCKLNLVDLAGSERODKTGATGDLKEGKIN 282
QY 283 RSLSITGLRVIAAALDMSGGQKKNOLVPRDYSVLTKLSDSGNSMTAMIAISPADIN 342
DB 283 LSLTALGNVISALVDCGSGH-----IPYRDSKLTLLQDSLGGNTKTVVNAVIGPADWN 336
QY 343 FEETLSLRVADSARKIRKHAHVNDPNARMRELKELAEQRLSKLQSSGGGGG----A 398
DB 337 VDETMSTLRVANRAKNIQNKPKINEDPKDAMLQFOEIKKLKEQLAARAAGGGGTMP 396
QY 399 GSGSGVE-----ESYPPD-----TPLEKQIVSQIQPDATVKKMSKAEIVFQNLQ 443
DB 397 SGGSPTOKIVERTEEVDPDIDAIKQAQRAELEAKMKSDISTEALDKAREEAAAK-KQ 455
QY 444 SEKLYRDLNOTWEEKLAKTETIHKEREAALELGISIEK 482
DB 456 LQAITDDQGTQAKQAARDALAKQAEBARAIAQ-ATEK 493

RESULT 14
Leishmanin-related antigen - Leishmania chagasi (fragment)
C:Species: Leishmania chagasi
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C:Accession: A47334
R:Burns Jr., J.M.; Shreffler, W.G.; Benson, D.R.; Ghallib, H.W.; Badaro, R.; Reed, S.G.
Proc. Natl. Acad. Sci. U.S.A. 90, 775-779, 1993
A:Title: Molecular characterization of a kinesin-related antigen of Leishmania chagasi
A:Reference number: A47334; MUID:93133867
A:Accession: A47334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-955 <BUR>
A:Cross-references: GH:I07879; NID:9308884; PIDN:AAA29254.1; PID:9308885
A:Experimental source: MHOM/BR/82/BA-2, CI
A:Note: sequence extracted from NCBI backbone (NCBI:122864, NCBI:P.122865)
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; P-loop
F:13-398/Domain: kinesin motor domain homology <KMOT>
F:122-125/Region: nucleotide-binding motif A (P-loop)

Query Match 20.2%; Score 814; DB 2; Length 955;
Best Local Similarity 28.1%; Pred. No. 1.8e-36;
Matches 253; Conservative 125; Mismatches 292; Indels 230; Gaps 24;

QY 7 IKVVVVRFPNARE--IDRGAKTIVMGNQITLTPP-----PGAEEKARKSGKTI 57
DB 13 VVSIVVRPLNERINAEPTGKTVYAAQAAAAYTVVKLGSSNNGSNAESKGTARRVAQD 72

QY 58 GPKAFAPDRSYNSF---DKNAPNARYQEDLPDGLVPLLDNAFKYNNCIFAYGQTSGK 114
DB 73 -----FOFDRVFSVETPDACGATPATQADVFRTIGYPLVQHAFGDFNSCLFAYGQTSGK 128

QY 115 SYSNMGK-----YGEHGVTPRCODMPEIRINELQKKNLCTVEVSYLEIYNERVDDL 168
DB 129 IYTMKGAADVSALESGNGVYPRICLIEIPARKSAVEAGHSRMIVELGYVEVYNERVSDLL 188

QY 169 NPSTKG-----NLKVRHPSTGTPYVEDLAKLVRSFOETENLMDGKNKARTVAATNM 220
DB 189 CKRKKVKVGGGEEVYVDVREHPSRGVFLGQRLVGVGSLDVLVRLIEIGNVHTASTKM 248

```

QY 221 NETSSSHAVPTLT-----TQKHDETKMDTEKVAKISLVLDLAGSERATSTGATGABLK 276  
 Db 249 NDRSSSHAIIMLLREERTWTKSGEITHTAGSSRNMLVLDAGSERVAQSQVQCFK 308  
 QY 277 EGAEINRSLSLGRVIAALADMS--GQKKNLQVYROSIVTLLKDSLGNSHWAMTAA 335  
 Db 309 EATHIMLSLTLTGRVIDVLDAMATGAKAQAQSVAPPKRSKLTFLKDSLGNSKTFMTAT 368  
 QY 336 TSPADINFEETLSLRYAQSARKIHNAVYEDPNAMRELKELAQRLKSSGGGG 395  
 Db 369 VSPALNFEETLSLRYAQSARKIHNAVYEDPNAMRELKELAQRLKSSGGGG 421  
 QY 396 GGAGSGGPPVE-----SYPTDPLEK 417  
 Db 422 --AGDPAYVELKKLLESFAQKRAADQALEREHNOVQRELLRYATEAKSELES 479  
 QY 418 QIVSIOQ-----PDATVKKSKAEIVQL 441  
 Db 480 RAMALOEEMTATROADQMALNLKPEQARKELLKEMAKDAALSKVRRKDAEIA 539  
 QY 442 NQSEKLYRLDNOTWEELAKTEIHKREAALEBLGISTEKGFGVPSKEMPHLVNLSID 501  
 Db 540 SREKL-----ESTVAQLEREQREVALDAL-----OTHORKLOEALSESE 581  
 QY 502 -----DPLABLCLYHVKPGOTRVGNVNDTQAEIRLNGSKILKECHCTFENVNVTIV 555  
 Db 582 RTAAERQDLQO--LITELASERTOLSOVVDRELTR--DLRIQTEYGETELARDVALCA 638  
 QY 556 PNE-----KAAVNGVNDKPRULSGSYRIILGDHPHFRNPEARERQEOSLHRS 610  
 Db 639 AQEMEARVHANFHLQTLLEATEDALR-----ERALAEERDAAR-- 682  
 QY 611 VTNQLGSPAGRHDTLSKASDAGDSRSDPLPHFGKDSQWFAVBAASALGLID 670  
 Db 683 -----LDAASTSON-----ARESACERLTSLE 705  
 QY 671 QKISHLTDDLDALFDVQKARVARGILVRDNEQSDSSQSPVPRDKYNTNSGNTIDNFSLD 730  
 Db 706 COLRAE--SEERAEALASQJATAAASASQADRENTATLEQRLS-----EARAEL 757  
 QY 731 AITWPGTFRSDDDGDLFFGDKSKQDAQSVN-----DVEELRQQAQOEALAKTAK 781  
 Db 758 ASQLEATAA-----KMSAQDRENTATLEQRLDSEERAEALASQLESTNAK 807  
 RESULT 15  
 B44259  
 Kinesin-related protein KIF3A - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 16-Jun-2000  
 C:Accession: B44259; S27872  
 J:Alzawa, H.; Sakine, Y.; Takemura, R.; Zhang, Z.; Nangaku, M.; Hirokawa, N.  
 J. Cell Biol. 119, 1287-1296, 1992  
 A:Title: Kinesin family in murine central nervous system.  
 A:Reference number: A44259; MUID:93077686  
 A:Accession: B44259  
 A:Molecule type: mRNA  
 A:Residues: 1-701 <A1Z>  
 A:Cross-references: EMBL:D12645; NID:g220469; PID:BAA02166.1; PID:g220470  
 A:Experimental source: brain  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118911)  
 C:Complex: heterodimer with KIF3B (PIR:A57107); the KIF3A/3B heterodimer associates with  
 C:Function:  
 A:Description: KIF3 complex is a motor protein that provides anterograde fast axonal tra  
 C:Superfamily: kinesin-related protein KIF3; kinesin motor domain homologous  
 F:Keywords: ATP; coiled coil; heterodimer; heterotrimer; microtubule binding; P-loop  
 F:15-351/Domain: head globular; #status predicted <ldg>  
 F:15-351/Domain: head globular; #status predicted <ldg>  
 F:100-107/Region: nucleotide-binding site N (P-loop)  
 F:369-599/Domain: helical rod #status predicted <ldg>  
 F:600-701/Domain: tail globular #status predicted <ldg>  
 F:105/Binding site: ATP (Lys) #status predicted

Query Match 19.9%; Score 800; Db 1; Length 701;  
 Best Local Similarity 31.88; Pred. No. 65e-36;  
 Matches 231; Conservative 110; Mismatches 231; Indels 154; Gaps 25;  
 QY 6 NIKVVVVRPNAREID---RGAKIVRMESNOTILTPPGAAEKARKSKGTIMDGPKAF 62  
 Db 14 NIKVVVVRPNAREID---RGAKIVRMESNOTILTPPGAAEKARKSKGTIMDGPKAF 62  
 QY 63 ADRSVNSDFKNAPYARQDFUDGLQVPLLDNAPKYNACIFAYQGTGSKSKSYMMGVG 122  
 Db 62 TFDTVF-----GPE--SQQLDYNILNARPIIDSVLEBYNGTIFAYQGTGTFITMGVYR 114  
 QY 123 KE---HGVIPRIQDMFRRELQKOKNLCTCTVEVSYLEIHYERYDLDLNPSTKGNLKYR 179  
 Db 115 AVPLGRVLPNFRSEFAHITGLHNAAGUTRF--LVNYSLEIHYERYDLDLNPSTKGNLKYR 172  
 QY 180 EHPSTGYPVEDLAKLVDFOSTFQENLMDGKNKARTVAATNNKETSRSRSHAFITLT--TQ 237  
 Db 173 ERPDGVYIKDLSAYVNNADMRLWTLGHKNSRGVATNNHSSRSHAFITITECSE 232  
 QY 238 KWHDEETKMDTEKVAKISLVLDLAGSERATSTGATGARKEGAEINRSLSLGRVIAALAD 297  
 Db 233 KGYDGNMNV--RMGKHLVLDLAGSERQAKTGATGORKLEATKINLSLSLGNVISALVD 289  
 QY 298 MSSGKKOKNQLVPRDYSVLTMLLKDSLGNSMTAMIAISPADINFEETLSLRYADSKA 357  
 Db 290 -----CKSTHVPYRNKSLTRLULDSLGNSKYMCMANIGPADYNYDETISTLRYANRAK 343  
 QY 358 RIKNHAVYEDPNAMRELKELAQRLKSSGG--GGGAGSGSGPPVE----- 407  
 Db 344 NIKKATREDPDMDALLQPKQKEIEELAKKLEEGEVSQDSISQSEDEEGELGDEGKE 403  
 QY 408 -----SYPTDPLEK-----IVS 421  
 Db 404 RKRREDQACKKYSPPKVNQENQAKIDEERKALETKLQWEEPRNKARAELEERKDLKA 463  
 QY 422 IQQPDATVKKSKAE-----IVEQLNQSEKLYRLDNOTWEELAKTEIHKREAA 472  
 Db 464 QEHQSULEKLSALEKKVIVGGVOLLAKAEQKLLFESNMELERRRRRAEQLKELKEK 523  
 QY 473 -----LEELGISTEKGFGVPSKEMPHLVNLSDDPLLABCLVYHVKPGOTRVGNVND 526  
 Db 524 EDERLDIEKYSLSQEEAOG--KTKKLKLYVTH-----LMAAKSEADLQOE 568  
 QY 527 TQAEIR--LNGSKI LKECHCTFEN--VDNV-----TIVPN-----EKAAYV 564  
 Db 569 HQRETEGELNQLRSRELALQMLIDANFPDYQRYMIENYVHNMDEIGENQLKCVYTG 628  
 QY 565 NGVYRDKPT---ELRSYRIILGDHPHFRNPEARERQEOSLHRSVITNSQLSGSNAP 621  
 Db 629 NNNKQPTPPDKKERDPEVDLS--HVV-LAYTEESL--RQSLMKLERPRTSKGAKRPT 683  
 QY 622 GRHDT 627  
 Db 684 GRRKRS 689

Search completed: Apr 11 25, 2001, 10:14:07  
 Job time: 283 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:13:34 ; Search time 26.31 Seconds  
(without alignments)  
1020.764 Million cell updates/sec

Title: US-09-235-416-1

Perfect score: 4030

Sequence: 1 MSGGKNKVVVRPFNARE.....ELRQOQMEALKTAQEP 784

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1668.5	41.4	1816	1 KFI8-HUMAN	O60333 homo sapien
2	1667	41.4	1103	1 KFI2-HUMAN	O43896 homo sapien
3	1663.5	41.3	1950	1 KFI4-HUMAN	Q12756 homo sapien
4	1663.5	41.3	1895	1 KFI1-MOUSE	P33173 mus musculus
5	1660.5	41.2	1816	1 KFI8-MOUSE	O60575 mus musculus
6	1591.5	39.5	1816	1 KFI1-RAT	P33176 rattus norv
7	1591.5	39.5	1884	1 U10-DEBL	O51576 caenorhabd
8	1584.5	38.6	689	1 KFI8-RAT	O68858 mus musculus
9	860	21.3	699	1 K122-STREP	O46872 strongyloce
10	848.5	21.1	747	1 KFI3-HUMAN	O15066 homo sapien
11	844.5	21.0	747	1 KFI3-MOUSE	O61771 mus musculus
12	827.5	20.5	742	1 K121-STREP	O46871 strongyloce
13	825.5	20.5	786	1 FL10-CHLRE	P46869 chlamydomon
14	814	20.2	955	1 KNL-LEICH	P46865 leishmania
15	800	19.9	701	1 KFI3-MOUSE	P28741 mus musculus
16	789.5	19.6	702	1 KFI3A-HUMAN	O9496 homo sapien
17	782.5	19.4	1231	1 KFI4A-MOUSE	P33174 mus musculus
18	772	19.2	928	1 KFIH-NEUCR	P48467 neurospora
19	770.5	19.1	796	1 KFI3C-HUMAN	O35066 mus musculus
20	769.5	19.1	796	1 KFI3C-RAT	O51565 rattus norv
21	765.5	19.0	793	1 KFI4-HUMAN	O4782 homo sapien
22	765.5	19.0	1232	1 KFI4A-MOUSE	O95239 homo sapien
23	743.5	18.2	935	1 KFI8-DEBL	P46867 drosophila
24	743.5	18.2	935	1 KFI8-HUMAN	O46873 syncephala
25	703	17.6	672	1 OSM3-CAEEL	P46973 homo sapien
26	691.5	17.2	963	1 KFIH-MOUSE	O61768 mus musculus
27	691	17.1	963	1 KFIH-STREP	P53978 strongyloce
28	687.5	17.1	1031	1 KFIH-MOUSE	O61768 mus musculus
29	684.5	17.0	1032	1 KFIH-STREP	Q12840 homo sapien
30	684	17.0	967	1 KFIH-LOLPE	P231613 loligo peal
31	676.5	16.8	1027	1 KFIH-MOUSE	P33175 mus musculus
32	673	16.7	1051	1 EG5-HUMAN	P52732 homo sapien
33	671.5	16.7	815	1 KFIH-CAEEL	P34540 caenorhabd

## SUMMARIES

RESULT	1	34	670.5	16.6	957	1	KF5C-HUMAN	O60282 homo sapien
KFI8-HUMAN	AC	35	669	16.6	956	1	KF5C-MOUSE	P28738 mus musculus
O60333	AC	36	668	16.6	2663	1	CENE-HUMAN	Q02224 homo sapien
01-OCT-2000 (Rel. 40, Created)	DT	37	649.5	16.1	975	1	KINH-DROME	P17210 drosophila
01-OCT-2000 (Rel. 40, Last sequence update)	DT	38	634	15.7	1056	1	K125-ARATH	P82266 arabidopsis
01-OCT-2000 (Rel. 40, Last annotation update)	DT	39	633	15.7	1184	1	BINC-EMENI	P17120 emericocella
KINESIN-LIKE PROTEIN KIF1B (KLP)	DE	40	631.5	15.7	883	1	YB3D-SCHPO	O14343 schizosacch
KIF1B OR KIAA0591	GN	41	630	15.6	1067	1	EG52-XENLA	Q91783 xenopus lae
Homo sapiens (Human)	OS	42	626.5	15.5	1111	1	KIP1-YEAST	P28742 saccharomyc
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	43	623	15.5	1060	1	EG51-XENLA	P28045 xenopus lae
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC	44	621.5	15.4	1066	1	K125-YOBRG	P28826 nicotiana t
K125-TOBRG	GN	45	610.5	15.1	1066	1	K125-DROME	P46865 drosophila
KIF1B_TaxID=9606;	GN							
SEQUENCE FROM N.A.	RP							
Bougueret L., Dufaur-Care I., Grel P.;	RA							
"DNA encoding a kinesin-like protein (hk1p) comprising biallelic	RT							
markers."	RT							
Patent number WO0063375, 26-OCT-2000.	RL							
[2]	RN							
SEQUENCE OF 479-1816 FROM N.A.	RP							
TISSUE=Brain;	RC							
MEDLINE=98290545; PubMed=9628581;	RX							
Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,	RA							
Nomura N., Ohara O.;	RA							
"Prediction of the coding sequences of unidentified human genes. IX.	RT							
The complete sequences of 100 new cDNA clones from brain which can	RT							
code for large proteins in vitro."	RT							
DNA Res. 5:31-39(1996).	RL							
SEQUENCE OF 1449-1816 FROM N.A.	RP							
Isoqul T., Ota T., Havaashi K., Sugiyama T., Otsuki T., Suzuki Y.,	RA							
Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,	RA							
Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,	RA							
Arata M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,	RA							
Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;	RA							
"NEO human cDNA sequencing project."	RT							
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	RL							
-!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A	CC							
MICROTUBULE PLUS END-DIRECTED MOTILITY (BY SIMILARITY).	CC							
-!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND	CC							
TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM (BY SIMILARITY).	CC							
-!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104	CC							
SUBFAMILY.	CC							
-!- SIMILARITY: CONTAINS 1 FHA DOMAIN.	CC							
-!- SIMILARITY: CONTAINS 1 PH DOMAIN.	CC							
This SWISS-PROT entry is copyright © It is produced through a collaboration	CC							
between the Swiss Institute of Bioinformatics and the EMBL outstation -	CC							
the European Bioinformatics Institute. There are no restrictions on its	CC							
use by non-profit institutions as long as its content is in no way	CC							
modified and this statement is not removed. Usage by and for commercial	CC							
entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).	CC							
or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .	CC							
EMBL: AX039604; CAC16629.1; .	DR							

DR EMBL; AB011163; BAA25517.1; -;  
 DR EMBL; AK022977; BAB14341.1; -;  
 DR InterPro: IPR000253; -;  
 DR Pfam: PF001849; -;  
 DR Pfam: PF00169; PH; 1;  
 DR Pfam: PF00498; FHA; 1;  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1;  
 DR PROSITE; PS00067; KINESIN\_MOTOR\_DOMAIN2; 1;  
 DR PROSITE; PS00006; FHA\_DOMAIN; 1;  
 DR PROSITE; PS00003; PH\_DOMAIN; 1;  
 KW Motor protein; Microtubules; ATP-binding; Coiled coil.  
 FT DOMAIN 1 361  
 FT MECHANOCHEMICAL (MOTOR).  
 FT COILED COIL (POTENTIAL).  
 FT COILED COIL (POTENTIAL).  
 FT COILED COIL (POTENTIAL).  
 FT COILED COIL (POTENTIAL).  
 FT COILED COIL (POTENTIAL).  
 FT COILED COIL (POTENTIAL).  
 FT PH.  
 FT NP\_BIND 1702 1799  
 FT ATP (POTENTIAL).  
 FT NP\_BIND 97 104  
 FT ATP (POTENTIAL).  
 SQ SEQUENCE 1816 AA: 204474 MW: AD62F0515978C783 CRC64;  
 Query Match 41.4%; Score 1668.5; DB 1; Length 1816;  
 Best Local Similarity 45.0%; Pred. No. 1.6e-79;  
 Matches 356; Conservative 126; Mismatches 172; Indels 137; Gaps 16;  
 QY 4 GNIKVVVRVPFNAFEDIRGAKVIRMEGNGQITLTPPGAEKARKSKITMDGPKAFA 63  
 DB 3 GASVYKVVVRVPFNSRSTSKESKIIQMGNSTSIINPKPE-----APKFS 51  
 QY 64 FDRSYWSP-DKNKPNYARQEDLFQDLGVPLDNPAGKYNNCIFAYGTGSKSYSMGYG 122  
 DB 52 FDIYSWHSFDPCCFQSNQRYNDIGKEMLLHAFESYNYVCIFAYGTGAGKSYTMGKQ 111  
 QY 123 KEH--GVIPRICQDMPRINELQKNJCTVEVSYLEYINERNVOLLNPSTKGNKVRE 180  
 DB 112 EESQAGIYQJCEELFEKIND--NCNEEYSYVENVSEIYECYRVDLLNPKNKNLVRRE 170  
 QY 181 HPSTGYPVEDLAKLVRFSEBIENLMDGKNKARTVAATNNNETSRSHAVFTLTKQHW 240  
 DB 171 HLLGTFVEDLKLAVTIDIALDMGNKARTVAATNNNETSRSHAVFTYFKKH 230  
 QY 241 DEETKMQTEKVAISLDAGSEATSTGATGARLKGAEINRSLTGLRVIAALDM-- 298  
 DB 231 DNETNLSTKVKSLIDVLGASRADSTGATGRKLEGANINKSLTGLRVISALAEVDN 290  
 QY 299 ---SSGOKKKQVLQVRSVLTWLLKSLGNSMTAMIAISPADINFEETLSTLYKADS 355  
 DB 291 CTSKSKKKTKDFPYRDSVLTWLLKSLGNSRTVAALSPADINFEETLSTLYRAD 350  
 QY 356 AKRTIKHAVNEDPNARMEKLEALQSKLOSSGGG-----GGAG----- 399  
 DB 351 AKQIKCNVAVNEDPNAKVLEKVEYRLKDLRAQGLDIDPLDIDTSGGSGKYLK 410  
 QY 400 -----GSGGPVEESYPTDPLEKQ-----IVSQIQPDATYKK 431  
 DB 411 DFQNNKHRYLLASQNRPGHFSTAGSKSLSS-PSGCSLSSQVGLTSTVSIQ--ERINST 467  
 QY 432 MSKAETVQLNQSKLYDLNQTWEELAKTETHKREAALEELGISTEK--CFVGPVPH 489  
 DB 468 PGGEAETLERKESKIIAELNETWEELKRTAEMERRALLAEMGVAIRDEGDTGVPS 527  
 QY 490 SKEMPHLYNDDPILLASLVYNIKPGQTCVNNVNDQTAETRLNKSILKEHCTFENV- 548  
 DB 528 PKKTHPLVNLNEDPLMSCLLYIKDITVQDAQRQDIVLSGAIHKECHIFPSR 587  
 QY 549 ----DNVYTVPNKAAVYNGVRIDKPLRSQVRIILGDPHFIFRNPPEAREPQRO 604  
 DB 588 SNSGEVITVLPCESETYVNGKRSQVQLSGNRLNGKHVFRNPNPEQAREK- 646  
 QY 605 SLLRHVNTNQLGSPAPGRDHRTLSKAGSDAGDSRSDPLPRGKQSDMFIYAREAS 664  
 DB 647 -----TPSETPTSEPVDTWPAQRELLE 668

QY 665 ALLGLDQK-----ISHLTDDELDAFDQ-----YOKARAVRGLVED 701  
 DB 669 K-QGIDMKQEMKRLQEMETLYKKEBADLLLEQQRDLVYESKQLQALQKOVETSLAET 727  
 QY 702 NEDSDSQSSSFP 712  
 DB 728 TEEEREEREVP 738  
 RESULT 2  
 KFC1C\_HUMAN STANDARD; PRT; 1103 AA.  
 AC 043896; O75186;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE KINESIN-LIKE PROTEIN KIF1C.  
 GN KIF1C OR KIAA0706.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=98352063; PubMed=9685376;  
 RA Borner C., Closse T., Mueller S., Moeller N.P.H., Ullrich A.,  
 RA Lammers R.;  
 RT "Characterization of KIF1C, a new kinesin-like protein involved in  
 RT vesicle transport from the Golgi apparatus to the endoplasmic  
 RT reticulum.";  
 RL J. Biol. Chem. 273:20267-20275(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K.-i., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 RA Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:169-176(1998).  
 CC -!- FUNCTION: MOTOR REQUIRED FOR THE RETROGRADE TRANSPORT OF GOLGI  
 CC VESICLES TO THE ENDOPLASMIC RETICULUM. HAS A MICROTUBULE PLUS END-  
 CC DIRECTED MOTILITY.  
 CC -!- SUBUNIT: MONOMER (POTENTIAL).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED, WITH MOST  
 CC ABUNDANT EXPRESSION IN HEART AND SKELETAL MUSCLE.  
 CC -!- PTM: PHOSPHORYLATED ON TYROSINE.  
 CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104  
 CC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U01329; AAC52117.1; -;  
 CC EMBL; AS024606; BA031681.1; -;  
 CC MIM: 603060;  
 CC InterPro: IPR000253; -;  
 CC InterPro: IPR001732; -;  
 CC Pfam: PF00498; FHA; 1;  
 CC Pfam: PF00232; Kinesin; 1;  
 CC Pfam: PF00360; Kinesin; 1;  
 CC PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1;  
 CC PROSITE; PS00467; KINESIN\_MOTOR\_DOMAIN2; 1;  
 CC PROSITE; PS00006; FHA\_DOMAIN; FALSE\_NEG.

KW Motor protein; Microtubules; ATP-binding; Coiled coil;  
 KW Phosphorylation.  
 KW DOMAIN 1 355 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
 FT DOMAIN 359 388 COILED COIL (POTENTIAL).  
 FT DOMAIN 438 479 COILED COIL (POTENTIAL).  
 FT DOMAIN 523 590 FHA.  
 FT DOMAIN 633 674 COILED COIL (POTENTIAL).  
 FT DOMAIN 828 872 COILED COIL (POTENTIAL).  
 FT NP\_BIND 97 104 ATP (POTENTIAL).  
 FT CONFLICT 669 689 Q -> R (IN REF. 2).  
 FT CONFLICT 955 962 LMGGGW -> SGGGGGL (IN REF. 2).  
 FT CONFLICT 976 977 NY -> XL (IN REF. 2).  
 SQ SEQUENCE 1103 AA; 123071 MW; 1148396DAB2EC10 CRC64;

Query Match 41.4%; Score 1667; DB 1; Length 1103;  
 Best Local Similarity 47.5%; Pred. No. 66-80;  
 Matches 354; Conservative 121; Mismatches 157; Indels 114; Gaps 16;

QY 4 GGNIKVVRVPFNAREIDRGAKCIVRMENQITLTPPGCAEKARKSGKTIIMDGPKAFA 63  
 DB 3 GASVKVAVRVPFNARETSQAKCVVSGNQTTSIINP-----KQSKDAPKSF 51  
 QY 64 FDRSYWFDKNAP--RYARQEDFDQGLVPLDIAFNKYNICIFAYGQTGSGKSYMMGY 122  
 DB 52 FDSYWSHSTEDPQFASQDQVYDIEGEMLIHAFEGYNVICIFAYGQTGAGKSYMMGRQ 111  
 QY 123 K--RHGVIPRIQDFERINELQKNIKTCVYSVLEITNERNVOLLNPSTGKLVRE 180  
 DB 112 EPGQGGVPLQEDBDFRSVNG -SAQLSVYSVSTREITCERNVOLLNPKRSGLR 170  
 QY 181 HPTSGPYVDLAKLVRSPOEITNMDENGNKARTVAATNNNTSSRSRAVFTLTQKWH 240  
 DB 171 HPILGPVQDLSKLVTSYADIDMDCGNKARTVAATNNNTSSRSRAVFTVTPQRCH 230  
 QY 241 DEETKMDTEKAKISLVLAGSERATSTGATCARLKEAGETNRSITSLRGTVAALADSS 300  
 DB 231 DQLTGLDSEKYSKISLVLAGSERADSSGARGMRLKGANINKSLTTLGKVISALADMS 290  
 QY 301 GQKKHQLVPRDVSVPVLLKLSIGNSMTAMTAISPADINFEETLSRLYADSARKIR 360  
 DB 291 -KKRSDFIPRDSVPLVLLKLNGLNSRPMIAAUSPADINFEETLSRLYADRTKQIR 349  
 QY 361 NHANYVEDPNAWRELKEELAQLSKASSGGGGGAGG-----400  
 DB 350 CHAIIEDPNARLRELQEVYARLELLMAQLSASALEGLTEGSGVGLPAYSSEPA 409  
 QY 401 -----SSGPEESYPPDTPLEQKIVNSIQPDATVKMKSAIVEQLNOSEKLVDI 451  
 DB 410 PVSPSPPTTHGELSPSPNT--ESQI-----GPEEMERLOETEKIAEL 454  
 QY 452 NOTWEKLAETEHKREAALEELGTIERK--GFQVPYISKEMPHLVNLSDDPLAELC 509  
 DB 455 NETWEKLRTELKREALLAEMGVAVREDDGTGVFGPKKTPHLVNLNEDPLNSEL 514  
 QY 510 VYTKPGQTRVGNVNDTPQAIRLNGSKILLKEHCTEN-----DNYVTYVPEKAAVY 564  
 DB 515 LYHIKGVTRKGVQDMD-----IKLTGQFIREOCLFRSIPQDGEVYVTEPECEGETY 570  
 QY 565 NGVRIDKTRILASGRIILGDFHIFRNHPEEARQROQSLLRHSYVNSOLGSPAPGR 624  
 DB 571 NGKLYTEPLVLSGNRIYKGNVYRPNHPEARQLRER-----GYPPP-----614  
 QY 625 DRTLKASGDACDSDSPSLPHFGKSDWYARERASAIIIGLQOKTSHLTDDELAL 684  
 DB 615 -----PGPSPVDPWNFAQKELLEQ--QGIDIKLE--MEKRLQDL 650  
 QY 685 FDDVQKAVRGLVEDNE---DSOS 707  
 DB 651 ENQYRKEEADLLLEQOOLYADSDS 676

RESULT 3

KFIA\_HUMAN STANDARD; PRT; 1690 AA.  
 Q12756;  
 AC 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE KINESIN-LIKE PROTEIN KFIA (AXONAL TRANSPORTER OF SYNAPTIC VESICLES).  
 GN KFIA OR ATSV.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid:9606;  
 RN 1; Nucleotide sequence; Coding sequence;  
 RS SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 EX MEDLINE=96299637; PubMed=8661001;  
 RA Furlong R.A., Zhou C.Y., Ferguson-Smith M.A., Affara N.A.;  
 RT "Characterization of a kinesin-related gene ATSV, within the tuberous  
 sclerosis locus (TSC1) candidate region on chromosome 9q34.";  
 RL Genomics 33:421-429(1996).  
 CC -!- FUNCTION: MOTOR FOR ANTEROGRADE AXONAL TRANSPORT OF SYNAPTIC  
 CC VESICLE PRECURSORS (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the European Bioinformatics Institute and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X90840; CAAG2346.1; -;  
 CC HSSP; P17119; 3KAR.  
 CC MIM; 601255;  
 CC InterPro; IPR000253; -;  
 CC InterPro; IPR001752; -;  
 CC InterPro; IPR001849; -;  
 CC Pfam; PF00498; FHA; 1.  
 CC Pfam; PF00169; PH; 1.  
 CC Pfam; PF00360; KINESINHEAVY;  
 CC PROSITE; PS00360; KINESINHEAVY.  
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 CC PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.  
 CC PROSITE; PS50006; FHA DOMAIN; 1.  
 CC PROSITE; PS50003; PH DOMAIN; 1.  
 CC Motor protein; Microtubules; ATP-binding; Coiled coil.  
 KW DOMAIN 1 361 MECHANOCHEMICAL (MOTOR).  
 FT DOMAIN 366 383 COILED COIL (POTENTIAL).  
 FT DOMAIN 429 462 COILED COIL (POTENTIAL).  
 FT DOMAIN 516 572 FHA.  
 FT DOMAIN 622 681 COILED COIL (POTENTIAL).  
 FT DOMAIN 801 822 COILED COIL (POTENTIAL).  
 FT DOMAIN 1575 1673 PH.  
 FT NP\_BIND 97 104 ATP (POTENTIAL).  
 SQ SEQUENCE 1690 AA; 191083 MW; DBDEC784624FBAD CRC64;

Query Match 41.3%; Score 1663.5; DB 1; Length 1690;  
 Best Local Similarity 46.7%; Pred. No. 266-79;  
 Matches 351; Conservative 125; Mismatches 167; Indels 109; Gaps 15;

QY 4 GGNIKVVRVPFNAREIDRGAKCIVRMENQITLTPPGCAEKARKSGKTIIMDGPKAFA 63  
 DB 3 GASVKVAVRVPFNARETSQAKCVVSGNQTTSIINP-----KQSKDAPKSF 51  
 QY 64 FDRSYWFDKNAP--RYARQEDFDQGLVPLDIAFNKYNICIFAYGQTGSGKSYMMGY 120  
 DB 52 FDSYWSHSTEDPQFASQDQVYDIEGEMLIHAFEGYNVICIFAYGQTGAGKSYMMGR 109



[4] SEQUENCE FROM N.A. (ISOFORM 2).  
RC SPANINACAP,  
RL Nakagawa T., Hirokawa N.:  
RT "Identification and characterization of a new kinesin superfamily  
KT Kif19a-beta".  
CC Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
CL -! FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A  
CC MICROTUBULE PLUS END-DIRECTED MOTILITY.  
CC -! SUBUNIT: MONOMER.  
CC -! SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND  
CC TUBULULAR STRUCTURES WITHIN THE CYTOPLASM.  
CC -! ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/BETA AND 3;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -! SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN  
CC TISSUE (THE CEREBELLUM AND CEREBRUM) WITHIN A SINGLE  
CC TYPE OF NEURONAL CELL.  
CC -! SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104  
CC SUBFAMILY.  
CC -! SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -! SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to: licensed@isb-sib.ch).

EMBL: D17577; BAA04503.1; -.  
DR EMBL: AF090190; AAF06718.1; -.  
DR EMBL: AF131865; AAD39438.1; -.  
DR EMBL: AB023656; BAA75243.1; -.  
DR HSSP: P33176; 1BG2.  
DR MGD: MGI:108426; Kif1B.  
DR InterPro: IPR000253; -.  
DR InterPro: IPR001752; -.  
DR InterPro: IPR001849; -.  
DR Pfam: PF00498; Fha\_1.  
DR Pfam: PF00223; Kinesin; 1  
DR Prosite: PS00411; KINESIN\_WHEATY.  
DR Prosite: PS50067; KINESIN\_MOTOR\_DOMAIN; 1.  
DR Prosite: PS50067; KINESIN\_MOTOR\_DOMAIN; 1.  
DR Prosite: PS50066; Kina.D0MAIN; 1.  
DR Prosite: PS50003; PH\_DOMAIN; 1.  
KW Motor protein; Microtubules; ATP-binding; Coiled coil;  
KW Alternative splicing.  
XW Domain 1 361  
FT DOMAIN 1 361  
FT DOMAIN 365 386  
FT DOMAIN 470 502  
FT DOMAIN 556 612  
FT DOMAIN 668 737  
FT DOMAIN 841 869  
FT DOMAIN 1702 1759  
FT PH 1702 1759  
FT NP\_SAND 289 304  
FT VSNDLC 289 304  
FT VASPLIC 394 434  
FT VARSPLIC 707 1196

YESKLAQVQVRSIAETAEETESEESEVEEVPKHQPHFEFLA  
QFAEKWKHSIQTSLLDHLNGMAVYLKEENAVLSVELKKVK  
POFVLVLTISIPVPPLEHLDLMNEXTHEDRPFRPVVAIV  
ODIKMGTFHWSLDKLQRDLDMRWMDRAGEVASQAQDS  
ETNKTCGDPDFYDRFHFWKLVGSSPTFGHCNLERADTPSP  
TESTADSDITELADEQQAMDDEPDDEAFVDOTGSDAGTEEG  
SELFSDGHDPFDFTLVGRGVLPVLSNLIIYPVLIHRV  
ALVSEKGVEGFRLVAQVAIADDEADPDVSGIRGSQTAKI  
SFDNVEYFQSDFSAMTRTSGLSLELRIVEGGQSSEVIS  
PREVNRHNDLDSKGTLDGKKVMWEFGSEIEGNHLKIGSA  
FTFRVTYQASGLPIPEAFIFPCOFNFUHRIDEAFNPBEPLKN  
MGRGSLPGTHQNIAVETVESVPSVDTLKTKPIVFVEFGH  
KQGLKPLGKPLGKPLGKPLGKPLGKPLGKPLGKPLGKPLGK  
KQGLKPLGKPLGKPLGKPLGKPLGKPLGKPLGKPLGKPLGK  
-> T (IN ISOFORM 2 AND ISOFORM 3).  
-> T (IN ISOFORM 2 AND ISOFORM 3).  
TDRKIDMSGSGGGLKQFNQKHNKVIHQRFQNFSTA  
-> T (IN ISOFORM 2 AND ISOFORM 3).  
YESKLAQVQVRSIAETAEETESEESEVEEVPKHQPHFEFLA

```

FT FT VARSPLIC 1197 1816
FT FT CONFLICT 117 117
FT FT CONFLICT 520 523
FT FT CONFLICT 909 909
FT FT CONFLICT 1608 1609
FT FT CONFLICT 1612 1612
FT FT CONFLICT 1784 1784
FT FT CONFLICT 1784 1784
FT FT SEQUENCE 1816 AA; 204080 MW; E316EC295138E5DE CRC64;

Query Match 41.2%; Score 1660.5; DB 1; Length 1816;
Best Local Similarity 45.1%; Pred. No. 4.1e-79;
Matches 357; Conservative 127; Mismatches 170; Indels 137; Gaps 17;

QY 4 GNTKVVYVVPFNAREIDRGACIVRMGNETQILTTPPGAEEKARKSGKTMGDKRAFA 63
DB 3 GASYKVAVVRPNARETSRESKESKICIQMGNSTSIINPKNPKE-----APKSES 51

QY 64 FDRSYNSF-DKNAPYARQEDFDGLGVPLDINAFKGYNNCFAYGOTSGKSYSHMGYG 122
DB 52 FDSYNSHTSPEDPCFASQNRVYNDIGREMLHAFEGYNVCFAYGOTGAGKSYTHMGQ 111

QY 123 KEH-GVPIRCQDMFRINELQKDNKTVEVSYLEYIYNNERVRLDLPNAPSKGNLKVRE 180
DB 112 EESQAGIIPQICEELPEKIND-NCNDEMSYSVSEYMEITYCYRVDLLAPNKNKGLVRE 170

QY 181 HPTSGYVEDIALVVRFSQIEINLMDGKNKARTVAATNNNETSSRSHAVFTLLITOKNH 240
DB 171 HPLGYPVEDUSKLAVSYTDIADMDAGNKARTVAATNNNETSSRSHAVFTVIFOKQ 230

QY 241 DEETKMDTEKAKIISLDLAGSERATSTGARGLKEGAEINRSLSLTGLRWITALADM-- 298
DB 231 DEINLSTETKYSKISLDLAGSERADSTGARGLKEGAEINRSLSLTGLRWISALAEVDN 290

QY 299 ----SSGQKKNQIIPYRUSVLTWLEKSLGNSMTAMIAISPADINFEETLSLYAUS 355
DB 291 CTSKSKKKKTDPIPRUSVLTWLEKSLGNSMTAMIAISPADINFEETLSLYAUS 350

QY 356 AKRIKHAVNEDPNAMIRELKEALQALSKLSSGGG-----GGAGG---- 400
DB 351 AKQICNAVINEDPAKIVRELEAEVTLKQALGGLDIDIDPLIDDYSGSGKYLK 410

QY 401 -----SGGVEYSEPPDPLEKQ-----IYVIOQPDATYK 431
DB 411 DFQNNKHRYLASENORCFNSTASGSLTSS-PSGSLNSQVGLTGVTSIQ--ERIMST 467

QY 432 ASKAEIVQLOKSEKLYBOLNOMTWEKIAKTEPHKPREALEELGISIEK--GFTGVPYH 489
DB 468 PCGEATELRESEKITAELNETWEKILRYKTEALIMREALLAEMGYATREDGGTGLVES 527

QY 490 SKEMPHLVNLSDDPLALCLVYNIKPGOTRGVNVNDYQAEIRLNGSKILKEHCTFENV- 548
DB 528 PKKTPHVLNLEDPLASECLLYIKDITRVGQDAERODIVLSGAHKEHCLFRSER 587

QY 549 ----DNVTVIPNKAAMVNGVYRDKPRLRSQYRIILGDFHFRFNHPPEAREARQEO 604
DB 588 SNTGEVITVLEKCRSETTYNGRVNAHPVOLRSNGRIILMGKNHVNFRNHPPEAREAR- 646

QY 605 SLLRHSVYNSQLGSPAGRHDRITLSKAGSDADGDSKSDPLPHFRGKDSMDPYARREMAS 664
DB 647 -----TPSAETPPEVDWTFQARELLE 668

```

---

```

QY 665 ALLGLDQK-----ISHLTDDDELAL-----FDDVKARAVRGL-----VED 701
DB 669 K-QGIDMQEMKRLQEMTEILYKKEKEADLLLEQORLDYESKQLQALQOVETSLAAET 727
QY 702 NEQDSQSQSF 712
DB 728 TEEEEEVEEP 738

RESULT 6
KFID_RAT STANDARD: PRT: 1097 AA.
AC 035787;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DR KINESIN-LIKE PROTEIN KIFID.
GN KIFID.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR;
RA Rogers K.R., Griffin M., Brophy P.J.;
RT "The secretory epithelial cells of the choroid plexus employ a novel
RT kinesin-related protein.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SURFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR HSSP: AJ000696; CA04248.1; -
DR InterPro: IP000253; -
DR InterPro: IP001752; -
DR Pfam: PF00225; Kinesin; 1.
DR Pfam: PF00360; KINESIN;
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN;
DR PROSITE: PS00467; KINESIN MOTOR DOMAIN;
DR PROSITE: PS00006; FHA DOMAIN; FALSE_NEG.
DR MOTOR PROTEIN; Microtubules; ATP-binding; Coiled coil.
KW DOMAIN 358 320
KW DOMAIN 320 320
KW DOMAIN 320 320
KW DOMAIN 437 478
KW DOMAIN 478 478
KW DOMAIN 478 478
KW DOMAIN 478 478
KW DOMAIN 478 478
KW ATP BIND 96 103
KW SEQUENCE 1097 AA; 122333 MW; 8EF40B1C759BA5B CRC64;

```

---

```

Query Match 40.1%; Score 1617; DB 1; Length 1097;
Best Local Similarity 45.9%; Pred. No. 3.9e-77;
Matches 357; Conservative 129; Mismatches 184; Indels 108; Gaps 20;

QY 4 GNTKVVYVVPFNAREIDRGACIVRMGNETQILTTPPGAEEKARKSGKTMGDKRAFA 63
DB 3 GASYKVAVVRPNARETSRESKESKICIQMGNSTSIINPKNPKE-----APKSFML---KA-S 50
QY 64 FDRSYNSF-DKNAPYARQEDFDGLGVPLDINAFKGYNNCFAYGOTSGKSYSHMGYG 122
DB 51 FDSYNSHTSPEDPCFASQNRVYNDIGREMLHAFEGYNVCFAYGOTGAGKSYTHMGQ 110

```







```

QY 458 KLAKEETFWK--EREANALESGISTKGVG 486
      I::I::: I::I::I::
Db 449 LQRRESELHKAQDDQKILNKAIOKKLIV 480

RESULT 10
ID KF3B_HUMAN
ID KF3B_HUMAN STANDARD; PRT: 747 AA.
OL5066;
AC
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF3B (MICROTUBULE PLUS END-DIRECTED KINESIN
DE MOTOR 3B) (RH0048).
DE KIF3B OR KIAA0359.
GS Homo sapiens (Human).
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
ON
[1]
SEQUENCE FROM N.A.
RP
RC TISSUE=Brain;
RX MEDLINE=9743984; PubMed=9205841;
RA Nagase T, Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Nagase T, Ishikawa K.-I., Nakajima D., Ohira M., Ohara O.,
RP Prediction of the coding sequences of unidentified human genes. VII.
RP The complete sequences of 100 new cDNA clones from brain which can
RP code for large proteins in vitro."
RL DNA Res. 4:141-150(1997).
RC
CC -1- FUNCTION: MICROTUBULE-BASED ANTERO-DIRECTED TRANSLOCATOR FOR
CC MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC ACTIVITY IN VITRO (BY SIMILARITY).
CC
CC -1- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B (BY SIMILARITY).
CC
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC in Basel, Switzerland. There are no restrictions on the use
CC of the Bioinformatics Institute, as long as its name and logo are
CC not used for profit-making activities, and as long as the
CC modified and unmodified data are made available for commercial
CC activities, license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB002357; BAA20815.1;
CC DR
CC BSSP: P56536;
CC DR
CC MIM: 603754;
CC DR
CC InterPro: IPR001752;
CC DR
CC Pfam: PF00225; kinesin.1;
CC DR
CC PROSITE: PRO0380; KINESIN_MOTOR_DOMAIN.1;
CC DR
CC PRINTS: PS00411; KINESIN_MOTOR_DOMAIN.1;
CC DR
CC Motor protein: Microtubules; ATP-binding; Coiled coil; Neurone.
CC KW
CC FT DOMAIN 1 345 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
CC FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).
CC FT DOMAIN 580 747 GLOBULAR.
CC FT NP_BIND 96 103 ATP (POTENTIAL).
CC FT DOMAIN 386 393 POLY-GLY.
CC FT DOMAIN 394 406 POLY-GLU.
CC FT DOMAIN 723 730 POLY-SER.
CC FT DOMAIN 747 AA; 85125 MW; 97FA4573AFAB7023 CRC64;
CC SQ
CC
CC Query Match 21.1%; Score 848.5; DB:1; Length 747;
CC Best Local Similarity 32.8%; Pred. No. 2.8e-37;
CC Matches 253; Conservative 120; Mismatches 239; Indels 139; Gaps
CC 26;
CC
CC QY 6 NIKVVVRVFPNAREIDRGACTIVME---GQQTILTPPGAEKARKSGKTIDGPKAF 62
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC
CC Db 9 SVRVVRCRWNGKEAKAASVDKVDVYKLGQVSVKKPKGTAHEN-----PKTF 57
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC
CC QY 63 AFDRSY-IFSEDKNAPNTARQEDFDLDGVLDPDIAFGKYNCFIAYGOTGSGKSYNMGY 121
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

	Query Match	21.18%	Score 848.5	DB.1	Length 747	
	Best Local Similarity	32.18%	Pred. No. 2.8e-37			
	Matches 253	Conservative 120	Mismatches 259	Indels 139	Gaps 26	
QY	6	NIKVVVYRFFNAREIDRGAQCIVRNE	--GNOTLITLPPGGAEEKARKSKT	IMDGPKAF	62	
Db	9	SVRVYRQPMGRKEAKSAQTKV	LVGLGVSVKPKPGTAHNE	--SKFTG	57	
QY	63	AFDSY--NSEDKNAPYRQEDFDLQGLV	LDLNAPKGYNGNCFAYG	GTGCKSSKVSVM	121	
Db	58	FEADVDNN--	--AKQFLYDSTFPLVDVSLG	FNQFNCTIFAYG	TGTGKTYMDEI	109

```

QY 122 ---GKEGVIPRICQDMFRRINELQKDKNTCTVEVSYLEIYNERVOLLNPTSTKGNLAV 178
DB 110 RCDPEKRGVINSFDHITHISRSNOQYL---VRASYLEIYQREIRDLSSKQTKREL 166
QY 179 REHPSTGPGVEDLAKLVRSFQBIENLMDGKNTVAATNNETSSRSHAVFTLTQK 238
DB 167 KERPDGTGVVLDLSSFTKSVKEIEHVNVGNQNSVGCATNNHSSRSHAFVITI--- 223
QY 239 WHDEETKMDTE---KVAKISLVDLAGSERATSTGATCARLKEGAEINRSLSTGLRVIAAL 295
DB 224 --ECSEVGLDGENHIRVGLKLVLDLAGSERQAKTGAQGERLKEATKINLSLALCNVISA 282
QY 296 ADMSSGKKKQNLVPRVDSVLTWLKLSLGSNMTAMIAAISPADINFEETLSTLRAYS 355
DB 283 VD-----GKSTHPIPRDSKLTPLQDLSLGNKTVVNVANVPASYNVEETLTLYANR 336
QY 356 AKRIKHAVNVNEDPNARIMRELKEALQSLKSSQSGG-----GGAGGSGGPV 405
DB 337 AKNINKPRVNEPDKALLREFOBETARLKAOLEKRSIGRRKRKRGGSGGGSGGEE 396
QY 406 EESTPPTPLEKQIVSIQDPATVKKSKAEIYVEQ-----LQSEKLYRLDQNTWEK 458
DB 397 EEEGE-----EGSEEGDDKDDYWRNSQOEKLETKRAIVEDHSLVAEEKM-----RLKREK 446
QY 459 LAKTEEIHKEAREAALEELGISIE---KGFGVPY---HSEKMPHLVNLSDPPLAECIV 510
DB 447 EKMMEDLREKDA--EMLGAKIKAMESKLLVGNKIVDHTNEQOKILEQKQRIEABQ--- 502
QY 511 YNKPQTRVGNVNDQTAERLNGSKILKEHCTFENVDNVYTVPNKEAAVYVNGVRID 570
DB 503 -----KRREIQQ--QMSRODETLEKE--TVSSLOQEVDI-----KTK 539
QY 571 KPTRLSGYRIILGDFHIFRPNHPEEAERQEQS-----LLRHSVNSQLSGSPAGRH 624
DB 540 KAKLEKQKVAETALHQBHEHKEHQEQLQETONELTRKLKHLTIENFI--PLEEK- 596
QY 625 DRTLSKAGSDAGDSKSDSPPLHFRGKD-----SDWYAR-----R 660
DB 597 SKINARAFDEEDHUKHPTITLKNQOMKRPVSAGYKRPLSQIARMSWIRPEARV 656
QY 661 EASAILGLDQKLSHLDDELADFDVQKARVROGLVEDNEDSQSSQF 711
DB 657 AENIVLLELDPSTRTDYGPAIPKQV--AALDAALQDEDEIQVNDASSF 705

RESULT 11
KIF3B MOUSE
ID KIF3B MOUSE STANDARD; PRT; 747 AA.
OC Q6173;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KIF3B (MICROTUBULE PLUS END-DIRECTED KINESIN MOTOR 3B).
GN KIF3B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_Taxid=10090;
RN (1).
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=96032268; PubMed=7559760;
RA Yamazaki H.; Nakata T.; Okada Y.; Hirokawa N.;
RA "KIF3B," a heterodimeric kinesin superfamily protein that works as a
RT microtubule plus end-directed motor for membrane organelle
RT transport;
RL J. Cell Biol. 130:1387-1399(1995).
CC -/- FUNCTION: MICROTUBULE-BASED ANTIGRADE TRANSLOCATOR FOR
CC MEMBRANOUS ORGANELLES. PLUS END-DIRECTED MICROTUBULE SLIDING
CC ACTIVITY IN VITRO.
CC -/- SUBUNIT: HETERODIMER OF KIF3A AND KIF3B.

```

```

CC -/- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstations
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL: D26077; BAA05070.1; -
CC HSP: P56536; 2KIN.
CC MGD: MGI:107688; Kif3b.
CC InterPro: IPR001752; -.
CC Pfam: PF00225; Kinesin; 1.
CC PRINTS: PR00380; KINESINHEAVY.
CC PROSITE: PS00411; KINESIN MOTOR_DOMAIN1; 1.
CC PROSITE: PS00067; KINESIN MOTOR_DOMAIN2; 1.
CC KW Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.
CC DOMAIN 1 345 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
CC FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).
CC FT DOMAIN 580 747 GLOBULAR.
CC FT NP_BIND 96 103 ATP (POTENTIAL).
CC FT DOMAIN 386 393 POLY-GLY.
CC FT DOMAIN 394 405 POLY-GLU.
CC FT DOMAIN 723 730 POLY-SER.
CC SEQUENCE 747 AA; 85288 MW; PA369A4190ECB47 CRC64;

```

```

Query Match 21.0%; Score 844.5; DB 1; Length 747;
Best Local Similarity 32.9%; Pred. No. 4.5e-37;
Matches 256; Conservative 116; Mismatches 253; Indels 153; Gaps 27;

```

```

QY 6 NIKVVVVRVFPNARE---IDRGAKCIVRMEGNQITLTPPGAEKARKSGKTIHQGPKA 61
DB 9 SVRVVVRPMNGKKAASDKVDVDVKL--GQVSVKNPKG-----TSHEPKT 56
QY 62 FAFDRSY-WSEFDKNAPYARQEDFDQDLYPLDNPAPKYNCFAYGQTSKGSXSMAG 120
DB 57 FTFDAYDWN-----AKQFLEDETFPLVDSVLQGFNGTIFAYGQTTGKVTMBG 108
QY 121 Y---GREGVITPRICQDMFRRINELQKDKNTCTVEVSYLEIYNERVOLLNPTSTKGNL 177
DB 109 VRGDPKRGVINSFDHITHISRSNOQYL---VRASYLEIYQREIRDLSSKQTKREL 165
QY 178 VRHPSTGPGVEDLAKLVRSFQBIENLMDGKNTVAATNNETSSRSHAVFTLTQK 237
DB 166 LKERPDGTGVVLDLSSFTKSVKEIEHVNVGNQNSVGCATNNHSSRSHAFVITI--- 223
QY 238 KWHDEETKMDTE---KVAKISLVDLAGSERATSTGATCARLKEGAEINRSLSTGLRVIAA 294
DB 224 --ECSEVGLDGENHIRVGLKLVLDLAGSERQAKTGAQGERLKEATKINLSLALGNVISA 281
QY 295 LADMSSGKKKQNLVPRVDSVLTWLKLSLGSNMTAMIAAISPADINFEETLSTLRAYD 354
DB 282 LVD-----GKSTHPIPRDSKLTPLQDLSLGNKTVVNVANVPASYNVEETLTLYAN 335
QY 355 SAKRIKHAVNVNEDPNARIMRELKEALQSLKSSQSGG-----GGGAGGSGGP 404
DB 336 RAKNINKPRVNEPDKALLREFOBETARLKAOLEKRSIGRRKRKRGGSGGGSGGEE 395
QY 405 VE---ESTYPTPLEKQIVSIQDPATVKKSKAEIYVE---OLNQSEKLYRLDQNTWEK 459
DB 396 EEEEGEEDGDDKDDYWRNSQOEKLEIKRA---IVEDHSLVAEEKM-----RLKREK 447
QY 460 AKTEEIHKEAREAALEELGISIE---KGFGVPY---HSEKMPHLVNLSDPPLAECIV 511
DB 448 KKMEDLREKDA--EMLGAKIKAMESKLLVGNKIVDHTNEQOKILEQKQRIEABQ--- 502
QY 512 NIKPQTRVGNVNDQTAERLNGSKILKEHCTFENVDNVYTVPNKEAAVYVNGVRIDK 571
DB 503 -----KRREIQQ--QMSRODETLEKE--TVYSIQQEVDI-----KTK 540

```



```

RT      J Cell Biol 126:175-188 (1994).
CC      -1- FUNCTION: PROBABLY INVOLVED IN FLAGELLAR ASSEMBLY AND MAINTENANCE.
CC      -1- MAY PLAY A ROLE IN FLAGELLAR SYNTHESIS.
CC      -1- TISSUE SPECIFICITY: FLAGELLAR AXONEME.
CC      -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC      II SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is not used in a
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch)
CC      -----
CC      EMBL: L33697; AAA21738.1; -.
CC      HSP: P17119; 3KAR.
CC      InterPro: IPR001752; -.
CC      Pfam: PF00225; kinesin.1.
CC      PRINTS: PS00380; KINESINHEAVY.
CC      PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC      PROSITE: PS00667; KINESIN_MOTOR_DOMAIN2; 1.
CC      Motor protein; Microtubules; ATP-binding; Coiled coil.
CC      DOMAIN 1 358 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
CC      FT DOMAIN 367 687
CC      FT DOMAIN 688 786
CC      FT NP_BIND 97 104
CC      FT DOMAIN 388 391
CC      FT DOMAIN 705 714
CC      FT DOMAIN 756 759
CC      FT DOMAIN 786 ASP.
CC      FT SEQUENCE 786 AA; 86671 MW; F90969203EB79F1B CRC64;
CC      SQ

```

[illegible]

```

RESULT_14
KINSLATCH
ID KINL LATCH STANDARD; PRT; 955 AA.
AC P46865;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE KINESIN-LIKE PROTEIN K39 (FRAGMENT).
DS KIN.
GS Leishmania chagasi.
OS Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxId=5669;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHON/BR/82 / ISOLATE BA-2;
RX MEDLINE=9313867; PubMed=8421715;
RA Burns J.M., Jr., Shreffler W.G., Benson D.R., Chalik H.W., Badaro R.,
RA Reed S.G.;
RT "Molecular characterization of a kinesin-related antigen of
RT Leishmania chagasi that detects specific antibody in African and
RT American visceral leishmaniasis.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).
CC C -! DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.
CC C -! SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC C
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licens@sib-sib.ch).
CC C
CC EMBL; LO7879; AAA29254.1; -.
DR HSSP; P17119; KNAR.
DR InterPro; IPR001752; -.
DR Pfam; PF00225; Kinesin_1.
DR PROSITE; PS0041; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
DR KQ Protein; Microtubules; ATPases; Coiled coil; Repeat.
FT DOMAIN 1 398
FT COILED COIL (POTENTIAL).
FT NP_BIND 426 >955
FT FT BIND 122 129
FT DOMAIN 704 >955
FT REPEAT 704 742
FT REPEAT 743 781
FT REPEAT 782 820
FT REPEAT 821 859
FT REPEAT 860 898
FT REPEAT 899 937
FT REPEAT 938 >955
FT NON_TER 955 955
SQ SEQUENCE 955 AA; 106168 MW; 8CA76815BE84C6E9 CRC64;

```

Query Match	20.2%	Score 814	DB 1	Length 955
Best Local Similarity	28.1%	Ned. No. 2.5e-35		
Matches 233	Conservative 125	Mismatches 392	Indels 230	Gaps 24
QY 7	IKVVRVRFNARE--IDRGAKCIVMREGNOTILTPP-----PQAEKARKSKTKTMD 57			
DB 13	VKYSVRPLNRENNAPETKTVAAQAAAVTVVKVLGSSNGSAAESMGTRARVAQD 72			
QY 58	GPKAFAPDRYSWFSF---DKNAPNPAQEDLPQDLGVPLLDNAFGYNNCTFAYGTQSGK 114			
DB 73	---PQDFHFWFSVETPDACCATPATQADVFRTIGYLVQAHFADFNCLFAYGTQSGK 128			
QY 115	SYSNMG-----YKGEHGVIPRICQMPFRINLQKDKNITCTVEVSLIETYNVRDLL 168			
DB 129	TYTNGADVSAISGNGVGPTRICLETFAFKASVEAQGSRWITELGVGVYVNRVSDLL 188			





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model!

Run on: April 25, 2001, 10:12:54 ; Search time 77.5 seconds  
(without alignments)  
1185.690 Million cell updates/sec

Title: US-09-235-416-1  
Perfect score: 4030  
Sequence: 1 MGGCNKVVVRVFNARE.....ELROQDWEALKTAQDF 784  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 374700 seqs, 117207915 residues  
Total number of hits satisfying chosen parameters: 374700  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SP TREMBL15.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phage.\*  
10: sp-plant.\*  
11: sp-rodent.\*  
12: sp-unclassified.\*  
13: sp-vertebrate.\*  
14: sp-virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1666.5	41.4	1770	11 Q92119	Q92119 mus musculus
2	1665	41.3	1103	4 Q75186	Q75186 homo sapien
3	1660.5	41.2	1816	11 Q9R0B4	Q9R0B4 mus musculus
4	1652.5	41.0	1816	11 Q9WVE5	Q9WVE5 mus musculus
5	1635	40.6	1773	5 Q9V7T6	Q9V7T6 drosophila
6	1630.5	40.5	1671	5 Q9N8L1	Q9N8L1 drosophila
7	1617	40.1	1097	11 Q35787	Q35787 rattus norv
8	1590.5	39.6	1584	5 Q18778	Q18778 caenorhabdi
9	1554.5	38.6	689	11 Q88558	Q88558 rattus norv
10	1525.5	37.9	2205	5 Q9NG02	Q9NG02 dictyostell
11	1412	35.0	1826	4 Q9NQ78	Q9NQ78 homo sapien
12	1396.5	34.7	1921	5 Q01349	Q01349 drosophila
13	1386.5	34.7	1921	5 Q01349	Q01349 drosophila
14	1284.5	31.9	1576	5 Q20898	Q20898 caenorhabdi
15	1174.5	29.1	1648	4 Q15058	Q15058 homo sapien
16	1147	28.5	928	5 Q09597	Q09597 caenorhabdi
17	1128	28.0	1174	5 Q9VB25	Q9VB25 drosophila
18	1035	25.7	1121	5 Q9V1P4	Q9V1P4 drosophila
19	1035	25.7	1121	5 Q16866	Q16866 drosophila

Q18390 drosophila  
Q9u102 homo sapien  
Q9uxn9 homo sapien  
Q93478 xenopus lae  
Q91160 arabidopsis  
Q9vrk9 drosophila  
Q91xv6 arabidopsis  
Q9ny24 homo sapien  
Q9u122 leishmania  
Q98657 rattus norv  
Q99640 gallus gall  
Q90000 drosophila  
Q90000 drosophila  
Q91784 caenorhabdi  
Q9u921 tetranychu  
Q9u179 leishmania  
Q9u175 arabidopsis  
P78718 nectria hae  
Q91dn0 arabidopsis  
Q19633 caenorhabdi  
Q9u141 leishmania  
Q9y6v4 homo sapien  
P87199 ustilago ma  
Q94229 drosophila  
Q9sne3 arabidopsis  
Q9gx11 mus musculu

## ALIGNMENTS

```

RESULT 1
ID Q92119 PRELIMINARY: PRT; 1770 AA.
AC Q92119:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE KIF1B-BETA.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR MOUSE;
RA Hsokawa; Hirokawa N.;
RT KIF1B-beta;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023656; BAA75243.1; -.
DR HSSP; P33176; 1BG2.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR001752; -.
DR INTERPRO; IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR PROSITE; PS50006; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil;
SQ SEQUENCE 1770 AA; 196850 MW; E44228501672E19 CRC64;

```

Query Watch 41.4%; Score 1666.5; DB 11; Length 1770;  
Best Local Similarity 47.3%; Pred No. 1.7e-90;  
Matches 353; Conservative 125; Mismatches 174; Indels 95; Gaps 15;  
QY 4 GGNKVVVRVFNAREIDRGAKIVMEGNQITLTPPPGAEKARKSKTMDGPKAPA 63





Db	458	PGGEALERLKESEKILAEUNETWEKLRTEARERALLAEMGYATREDGTLGYFS	27
Qy	490	SKEMPHLYNLSDPLLAECLVYNIKPGQTVGNWQVQOARILKLGSKILKEHCTFENV	548
Db	528	PKTHPLVPEKAAWCLSLYIKRGIVRQADARQDQIVLSGAHKEKHLFRSE	587
Qy	549	----DNVYTVNPEKAAWVNCVTDKPTLBSGVYILIGDPIFRNPPEARAEQEO	604
Db	588	SNTEGVIIVTLPEKSESYTYNGKVAHPVQLASGRLNKGKHFNPPEARAEK	645
Qy	605	SLLRHSYVNSQLSGFAPORIDRTILSKAGSDADGDSRSDPLPHFRGKSDMFTARREAS	664
Db	647	-----TPSACTSPSPVDVMTFAQRELLE	668
Qy	665	ATILGIDQK-----ISHLTDDLEAL-----FDVQKARAVRGL-----VED	701
Db	669	K--GGDKMKEKRIQLQEMILYKKEKEADALLLQOORLDYESKLQALQROVETRSLAET	727
Qy	702	NEDSDSQSSFP	712
Db	728	TEEEEEEVPP	738
RESULT 4			
IT	Q9WV5	PRELIMINARY; PRT; 1816 AA.	
DR	Q9WV5		
DT	Q9WV5	01-NOV-1999 (TRENBLrel. 12, Created)	
DT	Q9WV5	01-NOV-1999 (TRENBLrel. 12, Last sequence update)	
DT	Q9WV5	01-OCT-2000 (TRENBLrel. 15, Last annotation update)	
GN	KIFIB	KIFIB MAJOR ISOFORM	
GN	KIFIB		
OS	NCBI	Mus musculus (Mouse),	
OS	Eukaryota	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OS	Mammalia	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OS	NCBI	TaxID=10090;	
RP	[1]	SEQUENCE FROM N.A.	
RP	STRAIN-C57BL/6J	TISSUE-BRAIN;	
RC	Conforti L.H., Buckmaster A., Tarleton A., Brown M.C., Lyon M.F.,		
RC	Perry V.H., Coleman M.P.;		
RA	"The major brain isoform of Kifib lacks the putative mitochondria-		
RA	binding domain";		
RL	Mamm. Genome 10:0-0(1999).		
RL	EMBL; AF111863; AAC39438.1; -		
DR	ISSF; K31101; RG2.		
DR	WU; MG11081.26; kifb.		
DR	INTERPRO; IPR000253;		
DR	INTERPRO; IPR001732;		
DR	INTERPRO; IPR001849;		
DR	PFAM; PF00169; PH1;		
DR	PFAM; PF00225; Kinesin1;		
DR	PFAM; PF00498; FHA; 1.		
DR	PRINTS; PR00380; KINESINHEAVY.		
DR	PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.		
DR	PROSITE; PS50003; PH_DOMAIN; 1.		
DR	PROSITE; PS50006; FHA_DOMAIN; 1.		
DR	PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.		
KW	Motor protein; Microtubules; ATP-binding; Coiled coil.		
SO	SEQUENCE	1816 AA; 204153 MW; 95C9196A9DB4895A CRC64;	
Query Match 41.0%; Score 1652.5; DB 11; Length 1816;			
Best Local Similarity 45.0%; Pred. No. 1.2e-49;			
Matches 356; Conservative 127; Mismatches 171; Indels 137; Gaps 17;			
Qy	4	GQTKVWVPEVPEAREIDRACGVIRVNEGNQTLTPPPGAEKARKSKTKTMDQKAPA	63
Db	3	GASKYAVAVPEPNSGTSKESCIIQNGSTSLINFAKPE-----APAFS	51
Qy	64	FDSYVNSF--DNNAFWARQEDFDQDGLVPLDIAAFKGYNCIFAYGQCSGKSYWGVG	122



```

OY 304 --KKNOLVYRDSVLTMLKDSLGNSKTAIAATSPADINFEETLSTLYADRAKRIK 361
D 323 NTKKADPIFYRDSALTMLLRNLGNSTKTAIAATSPADINFEETLSTLYADRAKQVC 382
OY 362 HAVYNEDPNARMIRELKEELAQRLSKLOSSGGGGAGGSGPVESYPPDTPLEKQIVS 421
D 383 KAVYNEDANAKLIRELKEETQKLDLKAEG-----IEVEQSGPK-----VV 425
OY 422 IOODPATYKMSKAEI-----VQLONSEKLYRDLNOMTEKLAETE 463
D 426 CEKRDANKDELTVSTVPSKTSRNNSTTEMAVQLOQASEKLTAELNETWEKLRTE 485
OY 464 EHKREAAALBELGIST-EKGF-VGPHYSKEMPHLVNLSDDPLLAECLYVNIKPGQTRVG 521
D 486 EIRVOREAVFANGVAKEDGTVGVFSKPKPHLVNLDNPNLSECLLYIKESGTRLG 545
OY 522 NYNQQOQAEIRLNGSKILKEHCTFENOVNVTIVPNEKAAVMVNGVRIDKPTRLRSGYRI 581
D 546 THEANVPODIOQLSGSHILKEHCTFENKNSVTYLLPHKDAITYYNGRKLVEPVLKTSGRV 605
OY 582 ILGDFHIFRHNHPEARERQOSLLRHSVYNSQLSGSPAGRGHRTLSKAGSDADODSRS 641
D 606 IGRNHFRTNPEQARE-----LRDKI-----ETENEANEVE---KT 641
OY 642 DPLPHERKDSWIFYARREASAILGLD-----QKISHLTD-----ELDALFDVQ 689
D 642 DT-----QVDNFWAQCELEK-OGIDLKAEKMLNDLNEQYKREKLOADOQFEOR 735
OY 690 K--AR-AVRGLVDENEDSDSOSFFPVKYSNGTIDNFSLOTATMP----- 753
D 694 KTYEARIDALOKOVEOSMTSMYSYSFEDPHOEDVYTNPMWESCWTRAGLAANAF 753
OY 736 -----GTPRSDDDGDLFGDKKSKODASNDVDELRQOQAQ 772
D 754 RKNRYHOFTSLRDLGNALFL-----KEANAISVELKKYQFO 792

RESULT 6
QNNBL: PRELIMINARY: PRT: 1671 AA.
AC QNNBL:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KINESIN SUPERFAMILY MEMBER DUNC104.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Grossberger R., Saxton W.M., Dickson B.J.;
RT Characterization of the Drosophila unc104/RIF1A homolog, Dunc104.
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RS 561; AF247161; AAF419211;
SQ SEQUENCE 1671 AA; 169430 MW; EAC3FD3022360AF CRC64;

```

Query Match 40.5%; Score 1630.5; DB 5; Length 1671;  
 Best Local Similarity 44.3%; Pred. No. 2,1e-88;  
 Matches 364; Conservative 141; Mismatches 188; Indels 129; Gaps 20;

```

OY 6 NTKVVRVVRPNFARNDRGACIVRMENQOTILTPPGAEBEKARSGKTIHMGPKAFAD 65
D 3 SVKAVVRVVRPNFARNDRGACIEMAGATTATNP-----KVPNTDSYKRFNFD 54
OY 66 RSYNSFKNAPYARQEDLDGVLPLDNFAKGYNNICTEAFYQGTGSKSYNMGYGKE- 124
D 55 YSYNSHHDHADFQSKMYNDIGPMLQHSFDSYNCVFAYQGTGAGKSYTMNGQSEQ 114
OY 125 -HGVIPTICQDFRINELQKDKMLCTVTEVSYLEIYNERNVROLLNPSTGNLAKVREHPS 183

```

```

D 115 QEGIIPIKCDLFTRIQOTETD-DLKYSVEVSYMEICYERVDRDLNPKNGNLVRREHL 173
OY 184 TGYVEDIAKLQVFSQFIENLMDGSKARTVAATNNKETSRSRHAVFTLTLTQKHDEE 243
D 174 LGSYVEDLSKLAVTDYODIHDLDGSKARTVAATNNKETSRSRHAVFTIFFTORIDLM 233
OY 244 TKMDTEKYAKISLVLDLAGSRATSTGATGAKREGAETNRSIUSLTGRVYTAALADMSGK 303
D 234 TNUPTKVKYKISLVLDLAGSRADSTGAKREGANINAKSUTLAGYVISAIAESAKSK 293
OY 304 ---KKNOLVYRDSVLTMLKDSLGNSKTAIAATSPADINFEETLSTLYADRAKRIK 360
D 294 NTKKADPIFYRDSALTMLLRNLGNSTKTAIAATSPADINFEETLSTLYADRAKQIV 353
OY 361 HAVYNEDPNARMIRELKEELAQRLSKLOSSGGGGAGGSGPVESYPPDTPLEKQIV 420
D 354 KAVYNEDANAKLIRELKEETQKLDLKAEG-----I 386
OY 421 STQOPD-----ATVKMSKAE-----IVEQLONSEKLYRDLNOMTEKLAETE 466
D 387 EVDDEELTKSTVPSKTSRNNSTTEMAVQLOQASEKLTAELNETWEKLRTEIR 446
OY 467 KEREAAALBELGIST-EKGF-VGPHYSKEMPHLVNLSDDPLLAECLYVNIKPGQTRVGNV 524
D 447 VOREAVFANGVAKEDGTVGVFSKPKPHLVNLDNPNLSECLLYIKESGTRGTHG 506
OY 525 OTFOAETRLNGSKILKEHCTFENOVNVTIVPNEKAAVMVNGVRIDKPTRLRSGYRI 584
D 507 ANVPDIOQLSGSHILKEHCTFENKNSVTYLLPHKDAITYYNGRKLVEPVLKTSGRVILG 566
OY 585 DEHIFRHNHPEARERQOSLLRHSVYNSQLSGSPAGRGHRTLSKAGSDADODSRS 644
D 567 KKNHFRTNPEQARE-----LRDKI-----ETENEANEVE---KDT- 601
OY 645 LPHIFKDSWIFYARREASAILGLD-----QKISHLTD-----ELDALFDVQK- 690
D 602 -----QVDNFWAQCELEK-OGIDLKAEKMLNDLNEQYKREKLOADOQFEOR 654
OY 691 AR-AVRGLVDENEDSDSOSFFPVKYSNGTIDNFSLOTATMP----- 735
D 655 ENRDLRLOKOVEOSMTSMYSYSFEDPHOEDVYTNPMWESCWTRAGLAANAFRW 714
OY 736 -----GTPRSDDDGDLFGDKKSKODASNDVDELRQOQAQ 772
D 715 RYHOFTSLRDLGNALFL-----KEANAISVELKKYQFO 750

RESULT 7
O35787
ID O35787 PRELIMINARY: PRT: 1097 AA.
AC O35787;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE KINESIN-RELATED PROTEIN.
GN KIFID.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Perissodactyla; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR;
RA Rogers K.R., Griffin M., Brophy P.J.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ000696; CA04248.1;
DR HSSP: P17119; 3KAR.
DR INTERPRO: IPR000253;
DR INTERPRO: IPR001752;
DR PFAM: PF00425; kinesin 1.
DR PFAM: PF00498; FHA; 1.
DR PRINTS: P000380; KINESINHEAVY.

```



DB	112	ESBQAGITPTCCCELFPEKIND-NCNEDMSYSVS-SYMEITICERVVDLLPNKNGKINRVRE	169
QY	181	HBSTGPPVDLAKLVESFOETENMDGSKARTVAATNNKETSRSASHAVTLTJLTKQKH	240
DB	170	HPLLGPPVEDLSKLVATSYVDIADJMDAGNKARTVAATNNKETSRSASHAVTLTJLTKQK	229
QY	241	DEETKHOTEKVKAKISVLDLAGSERATSGTGAELKGEAEINRSLSTLGRVIAALADKGS	300
DB	230	DPETNLSTETKVKIKSVLDLAGSERADSGTGAELKGEAEINRSLSTLGRVIAALADKGS	289
QY	301	GROKKNLQVPTROSVTLMLKDSNGSMTAAITSPADINPEFTL--STLRVADSAGR	358
DB	290	-KKKTKDTFIPTRQSVTLMLKRENGSRTAWAALSPADINVDLTSTLSTLRVADRAGO	348
QY	359	IKNHAVYDENDPARNIRLEKLEALQKRSKQSGG-----	394
DB	349	ICKNVINDPNAKIVRELKVEVTRKLLQKQSGGL-----DIIDLNLQPNQKHRYLLASENQR	408
QY	395	-GGGAGGSGGPPVEEYPPDTPLEK-----IVSTQDPATVKYMSKAEIVQLNQSKELY	448
DB	409	PCNFSTASNGSLTSS-FSSCSLNSQAGLTSTVTSI--BRIMSTPGGEALERLKESEKII	465
QY	449	RLQNMWEKILAKTDEEIKEREAALDELGISIEK--GFVGGPYHKSMPHLVNLSDPPLA	506
DB	466	AELENETWEKILRTAIRMEREAALAEVGAIRDGSLTGTVSPKPTQPVNLNEDPLAS	525
QY	507	ECLVYNTKPGQTR-VGVNQDQTQAEIRLNGSKILK-EHCFTENVDN-----VVTYVFNK	559
DB	526	ECLLYIKDGTGFGQAQDERRQIVLSAGHTKKEHCIFRSENNKTGEIVTLPECR	585
QY	560	AAVMYNGVRDKTRLSRGYRIILDPHIFRNPHEBARQBOSLRHSVYNSQLSP	619
DB	586	STVYNGKVAHPVQLRSJGTVSSYKNVKNVFRNPFOARAEK-----	629
QY	620	AGRHEDTLKSGSDGSDGSRSPLPHFRGKDSQWVFARERASNLGLQOK	672
DB	630	-----TPSAETSPSPVMTFAQELLEK-QGIDMK	658
RESULT	10		
QNGQ2	QNGQ2		
AC	ID	PRELIMINARY;	PRT: 2205 AA.
AC	QNGQ2		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)		
DT	KINESIN UNCL04/KIF1A HOMOLOG.		
GN	UNCL04.		
GN	Dictyostelium discoideum (slime mold).		
OC	Bukaryota: Dictyostelidia; Dictyostelium.		
OK	NCBI_TaxId=44689;		
OK	NCBI_Accession=U00000;		
OK	SEQUENCE FROM N.A.		
OK	MEDLINE=20014990; PubMed=10545495;		
EX	Pollack N., de Hostos E.L., Turck C.W., Vale R.D.;		
RT	"Reconstitution of membrane transport powered by a novel dimeric		
RT	kinesin motor of the uncl04/kif1a family purified from		

J. Cell Biol.	147:493-506(1999).	[2]
SEQUENCE FROM N.A.		
Pollock N., Vale R.D.,		
Submitted (Mar-2000) to the ENBL/GenBank/DBJ databases.		
EMBL: AF243277; AF63384.1;		
SEQUENCE 2205 AA; 248001 AA; 02C5101E9D61C9ED CRC64;		
Query Match	37.9%;	Score 1525.5; DB 5;
Best Local Similarity	45.1%;	Pred No. 5' 6e-82;
Matches 333;	Conservative 118;	Mismatches 167;
Indels 121;	Gaps 13;	
6 NIKVVVRPFNAEIRDGAKCTVMEGNTLTPPGAEKARKSGKTINDPKAFAD 65		

```

Db 2 NVQVAVRVPENREKERNABELVQNNKSTILTRPSAL--BANPLAAPTAADKESFSD 59
QY 66 RSYWSFDKNAFYARDEDFODGLVPLDAPKAGYNNCIFAYGOTGSGSKSYSMWGYGKEH 125
Db 60 XSYWSYDSNDPHEASQSTVYNDLGRVKNADGKNSIFAYGOTGSGSKSYSMWGYGKEE 119
QY 126 GVIPRICODMFRRLNELQDKN--LTCVTVSVYLIYNERVRLDNLSTK--GNLKVYREH 181
Db 120 GLIPICEELFORLQSTSPNSNEQTIYKTVSVMEIYNEKVKDOLLNNKTKGLKVRNN 179
QY 182 PSTGPVVEDLAKLVRSFORBENLMDGNKARTVAATNMNETSSRSHAVFTLTJLQKWD 241
Db 180 PSTGPVVEDLKLAVKSESEIDLMDEGSKARTVASTNMNATSSRSHAVFTTIVTQSKID 239
QY 242 ETRMOTDEKVAKLSVLDLAGSERATSGATGARLKEGAEINRSLSLGRVIAALADSSG 301
Db 240 KTRGTALDRVSKLSVLDLAGSERANSTGATGVRLKEGANLNKSLTGKVISALAEINST- 298
QY 302 KQKNOLVPRDSVJLWLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSAKRINK 361
Db 299 -SKKAVFPYRDSVJLWLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSAKRINK 357
QY 362 HAVVNDPNAMRLRELKPELAQRSLKQSSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 421
Db 358 VAVVNDPAQSKLRELQGEVERLRAMDOGG-----QYHANDSKLMN----- 399
QY 422 IQOPDATYKMKSAEIVQLNOSEKLYRDLNQTWEKLAKEIHKEREAALBELGISTIE 481
Db 400 -SDVETVSTLN-----EKIDQELMAELNKSWEKLSAEAIREDRMAALKDGMVAIK 453
QY 482 KGFVGYSHKEMPHLVNLSDDPLLAECVLVYNIKPGQTRVGNVNOTQAEIRLNGSKILKE 541
Db 454 V-----VSSIPHLINLNDPLMSELSIYVYKGRTRIGRSDSEIPQDIIILNGIHK 506
QY 542 HCTFENVAVTVTVPN-----TKSPVDQIT-----MDYDFALNELAS-IQGTLAMSKII 557
Db 507 HCIFENAKVILSPNNFMNNNNKNSSTPTSSKSPKSEKENNDDDDGEG 566
QY 558 -EKAAMVYNGVRIDKPRLSGYRIILGDFHIFRNPPEARABEQSLRHSVNSQ 615
Db 567 KLDKSYIYNGVNEKPIILTGNRVILGNHIFRNPPEAIKATERNOTGGIVSS- 625
QY 616 LQSPAPGRHRTLSKAGSDADGSDSPHPGRKSDMVFYAREASAAILGDDQKISH 675
Db 626 -----TKSPVDQIT-----MDYDFALNELAS-IQGTLAMSKII 655
QY 676 LTD-----DELDAFDVQ 689
Db 656 INDKOYKKQMRALYDQIR 674

RESULT 11
QSN0T8
AC 001349 PRELIMINARY: PRT: 1826 AA.
ID 001349
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DE 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KINESIN-LIKE PROTEIN GAKIN.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RS 1
RA Kinesin-73, Lin L., Tibaldi E.V., Reinherz E.L., Chishti A.H.;
RA GAKIN, a novel kinesin-like protein associates with the human homolog
RA of Drosophila large tumor suppressor in T lymphocytes.;
RA J. Biol. Chem. 270:10720-10726 (1995).
RD EMBL: AF279865; AAF81263;
SQ SEQUENCE 1826 AA; 202665 MW; C614E7F3A89B89ED CRC64;

```

```

Query Match 35.08; Score 1412; DB 4; Length 1826;
Best Local Similarity 44.7%; Pred. No. 2,3e-75;
Matches 318; Conservative 114; Mismatches 193; Indels 86; Gaps 17;
QY 3 GSGNIKVVYVRPNAREIDRGAICTVMEGNOTILLTPPGAEKARKSKITMDGPKAF 62
Db 2 GDSKVVAVRIPMNRRETDLTFTCVVDYDANKVILNVPWNLNSGDAQGQ-----PKCF 56
QY 63 AFDRTSWSFDKNA--PYNARQEDLFDGLVPLDAPKAGYNNCIFAYGOTGSGSKSYSMWGY 121
Db 57 AYDHCWMSDSYKAYAGQDVIYKCLGENILQNAIDGYNACIFAYGOTGSGSKSYTMWCT 116
QY 122 KRGEGVLPICODMFRRLNELQDKN--LTCVTVSVYLIYNERVRLDNLSTK--GNLKV 178
Db 117 ADQPLGLPRLSGLPER--TQKEENEGSFKVEYVMEIYNEKVKDOLLDPKGSRLQAK 233
QY 179 REHPTGPVVEDLAKLVRSFORBENLMDGNKARTVAATNMNETSSRSHAVFTLTJLQK 278
Db 174 REHSLVGPYVDGLSKLAATSYKDIKESLSEGNKSKRTVAATNMNEESSRSHAVLTJLHT 233
QY 239 WHDEETKMDTEKVAKLSVLDLAGSERATSGATGARLKEGAEINRSLSLGRVIAALAD 298
Db 234 LYDAKSTSGEKGKLSVLDLAGSERATKTGAAGORLKEGSLNESLTLGLVISALADQ 293
QY 299 SSGKKNOLVPRDSVJLWLLKDSLGNSMTAMIAAISPADINFEETLSLRYADSAKR 358
Db 294 SAGK--NKNKFPYRDSVJLWLLKDSLGNSMTAMVATVSPAADNYDETSLTLRYADRAKH 352
QY 359 IKNHAVVNDPNAMRLRELKPELAQRSLKQSSGGGGGGGGGGGGGGGGGGGGGGGGGG 418
Db 353 IYNAVNVNDPNARIIRDLREVEKLRQLTKA----- 385
QY 419 IVSTQOPDATYKMKSAEIVQLNOSEKLYRDLNQTWEKLAKEIHKEREAALBELG 478
Db 386 -----EAMKSPDKLRESEKLTQEMTVTWEKLRTEIQAQROKQLESIGI 434
QY 479 STEKGFGVGYSHKEMPHLVNLSDDPLLAECVLVYNIKPGQTRVGNVNOTQAEIRLNGSK 538
Db 435 SLSQSS--GKYGKDDKDFVLNINADPALNELLVYLYKE--HTLIGSANQ--DIQCGMGI 488
QY 539 LKEHCTFE-NYDNNVTVIPNKAAMVNGVRIDKPRLSGYRIILGDFHIFRNPHE--EE 596
Db 489 LPHECHIDITSEGQVMLTPQKNTRTFVNGSSVSPILQHHGDRILWGNHIFRLNLPKKK 548
QY 597 ABAERQEQS---LLRHSVNSQLGSPAPGRHRTLSKAGSDADGSDSD--SPLPHFGRKD 652
Db 549 KKAEREDEDDQPSNKNNSSEQL-----DYDGSSESVSEYFN----- 587
QY 653 SDWFYARREAAASAILGDDQKISHLTDDELALFDVQKAVARGLVEDNE 703
Db 588 -NYTQAEVMTMKALGSDNPMOSIL-NSLEQHEEKEKRSALERQRIAYTHE 636

RESULT 12
QSN0T8
AC 001349 PRELIMINARY: PRT: 1921 AA.
ID 001349
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DE 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KINESIN-73.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyridioidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 1
RS 1
RA SEQUENCE FROM N.A.
RA MEDLINE=97188425; PubMed=9037010;
RA Li H.P., Liu Z.M., Nirenberg M.;
RA Kinesin-73 in the nervous system of Drosophila embryos.;
RA Proc. Natl. Acad. Sci. U.S.A. 94:10866-10871 (1997).
RA ENBL: U01788; AAB50404.1; -.

```







```

Db 555 YGGRALFLVAPNSSECCVNGKQITKTELLRNGRLVGNHFFVNCPC---KYNDMEQ 611
Qy 605 SILRHVSNTQLGSPAGRHDTLSKAGSDGSRSDSPHPKRGKDSDFYAREEAS 664
Db 612 SIMEDSTMFYD-----NDAHVNDANPI-----SS 637
Qy 665 ATLGDLQKISHLTDDDELDFDVOKARVVRGLVEDNEDSDS---QSSFFVRDKYMSNGT 722
Db 638 AY---DQWESVY-----LKHQEDKKAALQEQVEAFYIQSLT 673
Qy 723 IDNFSLOTATIMP-----TPRSDDGDDALFFGDKKSKQDASVNDVEELRQDQAAMEEALK 778
Db 674 AGGTFSTPMT-PGCLPTPTPTTGLPPPPFPANPKYSKFFYPWAQRKEEMFAESLK 732
Qy 779 TAKOE 783
Db 733 RLKAD 737

RESULT 15
ID O15058 PRELIMINARY: PRT: 1648 AA.
AC O15058;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE KIAA0042 PROTEIN.
GN KIAA0042.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RA 1.
RX 1.
RY 1.
SZ SEQUENCE FROM N.A. PubMed=7584044;
RA MEDLINE=66051398;
RA Nemura N, Nagase T, Miyajima N, Suzuki T, Tanaka A., Sato S.,
RA Seki N, Kawarabayashi Y, Ishikawa K., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 1:223-229 (1994).
DR ENBL: D26361; BAA05392.1;
DR HESP: P17119; 3KAR.
DR INTERPRO: IPR000253;
DR INTERPRO: IPR001752;
DR PFAM: PF00496; FHA. 1.
DR PFAM: PF00496; FHA. 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR PROSITE: PS0044; KINESIN_MOTOR_DOMAIN1. 1.
DR PROSITE: PS0047; KINESIN_MOTOR_DOMAIN2.
DR MACROBIO: M00000; ATP-binding. Coiled coil.
SQ SEQUENCE 1648 AA, 186490 MW, PDI4236687B7907 CRC64;

Query Match 29.1%; Score 1174.5; DB 4; Length 1648;
Best Local Similarity 34.8%; Pred. No. 2,6e-61;
Matches 281; Conservative 152; Mismatches 239; Indels 135; Gaps 17;

Qy 7 IKVVVRFPNAREIDRGAKIVRMGNTILTPPGAEKARKSGKTMGPKAFADR 66
Db 359 VTVAVRVRFETREKIEKASQVFMGKEITVEHP-----DTKOYIN---FIYDV 405
Qy 67 SYNSFKNPNVAREDFDLGVLGVLDDNFKYNNCIFYGQTSCKSKYSNMGYGKEHG 126
Db 405 SPNSFDECHPHASGTTVTEKLAANPLERAFGNTCLFYAGQTSCKSKYTMGMFSEEPG 465
Qy 127 VIPRICDWFNRINEKDKLNCVTVSVSLIYNEVRDL-----NPSTKGNLKVREH 181
Db 466 IIPRCEDLSQVARKQ-TQEVSYHLEASFEYNEKIHDLIVKCDENGQKQPLVYREH 524
Qy 182 PSTGCVEDLAKLVRSFOETNLNDEGNKARTVAATNNKNTSSSHSHTVLTLLQKWD 241

```

Search completed: April 25, 2001, 10:18:02  
Job time: 308 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: April 25, 2001, 10:12:41 ; search time 56.07 Seconds  
(without alignments)  
363.961 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_L\_357  
Perfect score: 1834  
Sequence: 1 MSGGNIKVVVRRFNARE.....PADINFEITSLTIRVADSAK 357  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 390729 seqs, 57163235 residues  
Total number of hits satisfying chosen parameters: 390729  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Genescs\_0401.\*
- 1: /SIDSI/gcgdata/genescq/genescq/AA1980.DAT.\*
  - 2: /SIDSI/gcgdata/genescq/genescq/AA1981.DAT.\*
  - 3: /SIDSI/gcgdata/genescq/genescq/AA1982.DAT.\*
  - 4: /SIDSI/gcgdata/genescq/genescq/AA1983.DAT.\*
  - 5: /SIDSI/gcgdata/genescq/genescq/AA1984.DAT.\*
  - 6: /SIDSI/gcgdata/genescq/genescq/AA1985.DAT.\*
  - 7: /SIDSI/gcgdata/genescq/genescq/AA1986.DAT.\*
  - 8: /SIDSI/gcgdata/genescq/genescq/AA1987.DAT.\*
  - 9: /SIDSI/gcgdata/genescq/genescq/AA1988.DAT.\*
  - 10: /SIDSI/gcgdata/genescq/genescq/AA1989.DAT.\*
  - 11: /SIDSI/gcgdata/genescq/genescq/AA1990.DAT.\*
  - 12: /SIDSI/gcgdata/genescq/genescq/AA1991.DAT.\*
  - 13: /SIDSI/gcgdata/genescq/genescq/AA1992.DAT.\*
  - 14: /SIDSI/gcgdata/genescq/genescq/AA1993.DAT.\*
  - 15: /SIDSI/gcgdata/genescq/genescq/AA1994.DAT.\*
  - 16: /SIDSI/gcgdata/genescq/genescq/AA1995.DAT.\*
  - 17: /SIDSI/gcgdata/genescq/genescq/AA1996.DAT.\*
  - 18: /SIDSI/gcgdata/genescq/genescq/AA1997.DAT.\*
  - 19: /SIDSI/gcgdata/genescq/genescq/AA1998.DAT.\*
  - 20: /SIDSI/gcgdata/genescq/genescq/AA1999.DAT.\*
  - 21: /SIDSI/gcgdata/genescq/genescq/AA2000.DAT.\*
  - 22: /SIDSI/gcgdata/genescq/genescq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	1834	100.0	Thermomyces lanugin
2	1119	61.0	Human Kinesin-like
3	1117	60.9	Human Kinesin-like
4	913.5	49.8	Gene 5 human secre
5	911.5	49.7	K39 polypeptide of
6	705.5	38.5	Leishmania chagasi
7	705.5	38.5	Amino acid sequenc
8	615.5	33.6	Drosophila kinesin
9	582	31.7	Drosophila kinesin
10	582	31.7	Drosophila kinesin
11	582	31.7	Drosophila kinesin

12	542	29.6	1269	21	Y77955
13	515	28.1	1518	21	Arabidopsis thalia
14	515	28.1	1662	21	Arabidopsis thalia
15	511.5	27.9	1460	21	Arabidopsis thalia
16	511.5	27.9	1462	21	Arabidopsis thalia
17	511.5	27.9	1604	21	Arabidopsis thalia
18	511.5	27.9	1606	21	Arabidopsis thalia
19	510	27.8	469	21	Arabidopsis thalia
20	495.5	27.0	452	21	Arabidopsis thalia
21	472	25.7	398	21	Arabidopsis thalia
22	464.5	25.3	324	19	Leishmania antigen
23	460	25.1	460	21	Human prostate can
24	433.5	23.6	979	20	Human Kinesin-rela
25	433.5	23.6	730	21	Xenopus laevis kin
26	433.5	23.6	1069	21	Arabidopsis thalia
27	433.5	23.6	1069	21	Arabidopsis thalia
28	433.5	23.6	1121	21	Arabidopsis thalia
29	424.5	23.1	829	21	Arabidopsis thalia
30	424.5	23.1	834	21	Arabidopsis thalia
31	404	22.0	790	21	Arabidopsis thalia
32	404	22.0	794	21	Arabidopsis thalia
33	404	22.0	814	21	Arabidopsis thalia
34	403	22.0	147	21	Human secreted pro
35	379.5	20.7	410	21	Human prostate can
36	324	17.7	959	21	Arabidopsis thalia
37	317	17.3	726	21	Arabidopsis thalia
38	302	16.5	154	21	Human ORFX ORF425
39	280	14.2	1201	20	Drosophila sp. Cos
40	211.5	11.5	243	21	Human cancer asoc
41	162	8.9	482	21	Human ORFX ORF1485
42	162	8.9	482	21	Human ORFX ORF1485
43	107	5.8	338	21	Arabidopsis thalia
44	107	5.8	338	21	Arabidopsis thalia
45	107	5.8	386	21	Arabidopsis thalia

ALIGNMENTS

RESULT	1
ID	Y06618 standard; Protein; 784 AA.
XX	XX Y06618;
XX	XX AC Y06618;
XX	XX 26-OCT-1999 (first entry)
XX	XX Thermomyces lanuginosus kinesin motor protein TL-gamma.
XX	XX TL-gamma; kinesin; motor protein; microtubule; unc-104; infection;
XX	XX neurodegenerative disease; Alzheimer's disease;
XX	XX Parkinson's disease; Huntington's disease;
XX	XX amyotrophic lateral sclerosis.
XX	XX Thermomyces lanuginosus.
XX	XX PN MO9937659-A1.
XX	XX PD 29-JUL-1999.
XX	XX PF 22-JAN-1999; 99WO-0501355.
XX	XX KR 23-JAN-1998; 98US-0072361.
XX	XX XX (RECC ) UNIV CALIFORNIA.
XX	XX PT Goldstein LSB, Sakowicz R;
XX	XX DR WPI; 1999-493950/41.
XX	XX DR N-PSDB; X87656.
XX	XX PT New nucleic acid encoding microtubule motor protein, used for
XX	XX diagnosis of fungal infection and neurodegenerative disease

```

XX PS Claim 5; Page 70-71; 75pp; English.
XX
CC This sequence represents Thermomyces lanuginosus TL-gamma, a novel
CC ATP-dependent, plus end-directed microtubule motor protein that is
CC a member of the unc-104 family and kinesin superfamily. The
CC invention provides TL-gamma nucleic acids (see X87656), proteins
CC and antibodies, and methods of screening for TL-gamma modulators
CC potentially useful for treating hyphal fungal infections and
CC diseases caused by mutated TL-gamma, e.g. neurodegeneration
CC involving anterograde axonal transport, such as Alzheimer's,
CC Parkinson's or Huntington's diseases or amyotrophic lateral
CC sclerosis. Detection of TL-gamma allows differentiation between
CC hyphal and non-hyphal fungal infections.
XX
SQ Sequence 784 AA:

Query Match 100.0%; Score 1834; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 9,3e-173;
Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 1 MSGGNIKVVVRPNAREIDRGAKCIVRMGNQITLTTPPGAEEKARKSGKTIMDGP 60
DB 1 ms99gnlkvvrvrpnareidrgakcivrmgnqtltppgaeeakarksgktimdgpk 60

QY 61 AFADRSYSDKNAPNAREQDLFDQGLVPLLDNAPKGNVNCIPAYGTGSGKSYSMWG 120
DB 61 afadrsysfdknapnareqdlfdqglvplldnarkgynncifaygtgsgksysmwy 120

QY 121 YGKHGVIPRICQDMFPRINELQKNLCTCTVEVSYLEIYNRVRDLLNPSFKGNLKVRE 180
DB 121 ykxhgvipricqdmfprinelqknlctctvevsyleiynrvrdllnpsckgnlkvre 180

QY 181 HPSTGYPVEDIAKLIVRSQEIENLMDGKNKARTVAATNNETSSRSNAVFTLTQKWH 240
DB 181 hpsgypvediaklivrsqeiennlmdgknkartaatnnetsrsnavftltqkwh 240

QY 241 DEETKMDTEKAKISIVOLAGSRATSGATGARKKGAENRSISLIGRVYTAALADMSS 300
DB 241 deetkmdtekakisiivolagsratstgatgarkkgaenrsisligrvytaalamss 300

QY 241 DEETKMDTEKAKISIVOLAGSRATSGATGARKKGAENRSISLIGRVYTAALADMSS 300
DB 241 deetkmdtekakisiivolagsratstgatgarkkgaenrsisligrvytaalamss 300

QY 301 GKQKKNQVLPYRDSVLTWLLKDSLGNSMTAMTAISPADINFEETLSIRYADSAS 357
DB 301 gkqkknqlpyrdsvltwllkdslgnsmtamtaiispadinfeetlsiryadsas 357

RESULT 2
ID B36227
XX B36227 standard; Protein; 1816 AA.
AC B36227;
XX
DT 19-FEB-2001 (first entry)
XX
DE Human kinesin-like protein HKLP SEQ ID NO: 4.
XX
KW Human; kinesin-like protein; HKLP; cell division; cancer;
KW intracellular transport; neurological disorder; infertility;
KW biallelic marker; spontaneous abortion; neonatal chromosome disorder;
KW aneuploidy.
OS Homo sapiens.
XX
PN WO200063375-A1.
XX
PD 26-OCT-2000.
XX
PF 20-APR-2000; 2000WO-IB00562.
XX
PR 20-APR-1999; 99US-0130217.
XX
PA (GENSET ) GENSET.

```

```

XX PI Bougueleret L, Dufaure-Gare I, Grel P;
XX
DR NPI; 2000-665242/64.
XX
DR N-PSDB; C66550.
XX
PT An isolated or purified human kinesin-like protein (HKLP) encoding
PT polynucleotide used to detect HKLP polynucleotides in a sample
PT comprises a contiguous span of at least 12 nucleotides.
XX
PS Claim 46; Page 189-192; 199pp; English.
XX
CC The present invention describes the coding and protein sequences of the
CC human kinesin-like protein HKLP. It is thought that the protein could be
CC involved in neurological disorders, infertility, spontaneous abortion,
CC neonatal chromosome disorders, aneuploidy and cancers. This is due to its
CC function in the movement of microtubules. The protein shows homology to
CC the murine KIF1A and KIF1B proteins. The sequences disclosed in the
CC invention can be used in the isolation of similar human proteins and in
CC vector production. In addition, the biallelic markers shown can be used
CC in disease diagnosis and population studies.
XX
SQ Sequence 1816 AA:

Query Match 61.0%; Score 1119; DB 21; Length 1816;
Best Local Similarity 60.8%; Pred. No. 2.9e-101;
Matches 220; Conservative 59; Mismatches 63; Indels 20; Gaps 5;

QY 4 GGNITKVVVRPNAREIDRGAKCIVRMGNQITLTTPPGAEEKARKSGKTIMDGP 63
DB 3 gasykvavrvrpnarsretskscilgmqgnstsiopknpke-----apksfs 51

QY 64 FDRSYNSF-DKNAPNTARQEDLFQDLGVPLLDNAPKGNVNCIPAYGTGSGKSYSMWG 122
DB 52 faysyhshtspedcfaeqnrvyndigkmlhahfegynvcifaygtgagksytmmgkq 111

QY 123 KEH--GVIPRICQDMFPRINELQKNLCTCTVEVSYLEIYNRVRDLLNPSFKGNLKVRE 180
DB 112 eesagqlpqlceelfekind-ncheemsyvsveymelycervrdllnpskgnlrvre 170

QY 181 HPSTGYPVEDIAKLIVRSQEIENLMDGKNKARTVAATNNETSSRSNAVFTLTQKWH 240
DB 171 hp1lgpyvedislavtsytdiadlmdagnkartaavaatnnetsrsnavftltqkwh 230

QY 241 DEETKMDTEKAKISIVOLAGSRATSGATGARKKGAENRSISLIGRVYTAALADM-- 298
DB 231 dnetnisteekvaksisiivolagsratstgatgarkkgaenrsisligrvytaalamd 290

QY 299 ---SSGQKKNQVLPYRDSVLTWLLKDSLGNSMTAMTAISPADINFEETLSIRYADS 355
DB 291 ctskskkskkskdcfipyrdsvltwllrenlgnmrtsamvaaispadinfeetlsiryad 350

QY 356 AK 357
DB 351 AK 352

RESULT 3
ID Y51328
XX Y51328 standard; Protein; 1103 AA.
AC Y51328;
XX
DT 17-APR-2000 (first entry)
XX
DE Human KLIMP protein.
XX
KW KLIMP; kinesin-like motor protein; cytosolic; anticonvulsant; human;
KW anti-Alzheimer; anti-Parkinsonian; antidiabetic; anti-ulcerative; cancer;
KW immunomodulatory; antiinflammatory; anti-AIDS; antirheumatic; treatment;
KW antiarthritic; diagnosis; neurologic disorder; vesicular transport.
XX

```





XX		10-JAN-1994;	94W0-USO0324.	
PP				
XX				
PR		15-JAN-1993;	93US-0006676.	
PA		(IASY-) TASY5 CORP.		
XX				
PI	Reed SG;			
XX				
DR	WPI, 1994-24&02/30.			
XX	NFSUBB; Q70152.			
PT	Diagnosis of Leishmaniasis - by determining the presence of			
XX	antibodies that bind to a K39 repeat unit antigen			
PS	Disclosure; Page 12-15; 20pp; English.			
XX	The K39 polypeptide comprises a number of repeated units (described			
CC	in R57366). Detection of antibodies directed against this repeated			
CC	unit in a patients sample is indicative of leishmaniasis. The			
CC	antigenic repeat unit can itself be used as a vaccine to protect			
CC	against infection by a leishmania parasite.			
SQ	Sequence 955 AA;			
		Query Match	38.5%; Score 705.5; DB 15; Length 955;	
		Best Local Similarity	42.7%; Pred. No. 1e-60;	
		Matches 163; Conservative 58; Mismatches 126; Indels 35; Gaps		
QY	7 IKVVRVPFPNARE--IDRGAKCTVMGNGQTITTP-----PGAEKARKSGKTIMD 57			
Db	13 vkvsrvrplnerennapegtkvcaagaavvtvkvlgsnmgaaesmgarrvagd 72			
OY	58 GPKAFAFDSTWSF---DKNPANYARQEDJFDJLGVPLDDNAPFYNCIFAYGTGGSK 114			
Db	9-----lqdnwlvavetpacgatpcagadvrtlgpylvqlvadfnscifaysgtggsk 128			
OY	115 SYSNWG-----YGEHCYIPRICOMPRIHELXOKNLCTCYEYLYETYNERYVDLL 168			
Db	129 tyumngadvsalgggvgtpcrlclelfarkvaeaghsrwlvclgyveyynervsdll 188			
QY	169 NPSTKG-----NLKVREHPSVGYPVEDLAKLVRSFQELDLMDGKNKARTVAATNN 220			
Db	189 grkrkgkvggeevydvdrhepsrgvfleqgrlvevgslddvrvllelgnvrrhtastkm 248			
QY	221 NETSSRSHAVFTLTL----TKMDEETEKMDTEKYAKLSLDLAGSERSTATSGTGARKLK 276			
Db	249 ndrsrhraimillireertumtksegetirtagsksrmmlvdlegservagsgvegqdkf 308			
QY	277 EGAEFNISLSTGRVIYALADNS-GCKKKNOIWPFRDSVLTWLLKSLGSGNKGNMTAA 335			
Db	309 eathinslttgivdladmmtktagdyaprfddkkltilkdslygnakflmfat 368			
QY	336 ISPADINPEELSTRVADSANK 357			
Db	369 vepsalnysaelstrvasrar 390			
RESULT	7			
ID	W03691			
AC	W03691 standard; Protein; 955 AA.			
XX	W03691;			
DT	09-MAR-1997 (first entry)			
XX	Leishmania chagasi K39 antigen.			
KW	Leishmania chagasi; acidic ribosomal antigen; LcpO;			
OS	eipote; K39.			
XX	Leishmania chagasi.			











```

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.1%; Score 515; DB 21; Length 1518;
Best Local Similarity 34.5%; Best No. 1.7e-41;
Matches 137; Conservative 63; Mismatches 119; Indels 78; Gaps 12;

OY 17 NAREIDRGAKC---IVRMGECNOTILTPPGAEKARKSGKTIDGPKAFAPDRSYSPD 72
DB 41 hnpdlrdntspdhrrmrnk-nplprppspnplkrksaetates----gfsds----- 90

```

```

OY 73 KNAPNYARQEDLPDGLVPIILDNAPKGYNNCTFAYGOTGSGKSYNMG-----VG 122
DB 91 -----gvkeqmfqlygplyvencisgfnssvfayqgtsqgktyumgpgangliehlcg 144
OY 123 KEHGVTPRTCDOMFRINE-----LOKDKNLTCTVEVSYLEINERVRDLN- 169
DB 145 dqrgltprferlifarikevmpgfvcfvvltlnkssmlkg-sitovgahyrsfswlby 203
OY 170 -----PSTKGNL-----KVREHPSTGPGYVEDLAKLVRSFOE 201
DB 204 lslddlqrannrptgpkpdkgdkbhvhrfslmcbqiredvksgyvveniteeyvknld 263
OY 202 IENLMDEGNKARTVAATNNNETSSRSHAVFTLTQKWHDEETKMDTEKVAISLVDLAG 261
DB 264 vslliklignrrtgatsvntessrhcvfctvvesrcknvadglsfksrlnlvdag 323
OY 262 SERATSTGATGARKEGAENRSLSTLGRVIAALADMS-SGOKKKNOLVPYSDSVLTWLL 320
DB 324 serqkstgaagerikeagninrslsglnlinlaeisqgkprh---ipyrdsrltfill 380
OY 321 KDSLGNSMTALAAISPADINFEETLSTLRYADSAK 357
DB 381 qeslgnnaklamvcavspqrs--etfstlrfraqrak 415

RESULT 14
G31282
ID G31282 standard; Protein; 1662 AA.
XX
AC G31282:
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37541.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126284.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128845.
PR 21-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130549.
PR 23-APR-1999; 99US-0130850.
PR 28-APR-1999; 99US-0131149.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

```

PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 22-MAY-1999; 99US-0135629.  
PR 23-MAY-1999; 99US-0136021.  
PR 24-MAY-1999; 99US-0136392.  
PR 25-MAY-1999; 99US-0136792.  
PR 26-MAY-1999; 99US-0137222.  
PR 27-MAY-1999; 99US-0137628.  
PR 03-JUN-1999; 99US-0137702.  
PR 04-JUN-1999; 99US-0137704.  
PR 05-JUN-1999; 99US-0138054.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 12-JUN-1999; 99US-0138687.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 19-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 19-JUL-1999; 99US-0142940.  
PR 19-JUL-1999; 99US-0143271.  
PR 13-JUL-1999; 99US-0143272.  
PR 14-JUL-1999; 99US-0143634.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145066.  
PR 21-JUL-1999; 99US-0145068.  
PR 22-JUL-1999; 99US-0145067.  
PR 22-JUL-1999; 99US-0145068.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147493.  
PR 08-AUG-1999; 99US-0147494.  
PR 10-AUG-1999; 99US-0148173.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151086.  
PR 27-AUG-1999; 99US-0151086.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157423.  
PR 06-OCT-1999; 99US-0157423.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160865.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.

```

PR   25-OCT-1999;          99US-0161406.
PR   26-OCT-1999;          99US-0161359.
PR   26-OCT-1999;          99US-0161360.
PR   26-OCT-1999;          99US-0161360.
PR   28-OCT-1999;          99US-0161920.
PR   28-OCT-1999;          99US-0161920.
PR   28-OCT-1999;          99US-0161920.
PR   28-OCT-1999;          99US-0161920.
PR   28-OCT-1999;          99US-0161933.
PR   29-OCT-1999;          99US-0162142.

Query Match      28 18; Score 515, DB 21; Length 1662;
Best Local Similarity 34.5%; Pred No. 2e-41;
Matches 137; Conservative 63; Mismatches 119; Indels 78; Gaps 12;

QY   17 NAREIDRGAKC---IVRMGNETLTITPPCAEKAARKSGKTMDCGRAPAFDRSWSFD 72
DB   41 nappldrntstpdhrsrmak-nplpprpnpnplrklsaetates----gfsds----- 90
QY   73 KNAPTYARQEDLDQDLGVPLDPAFNGYNCFAYCGTGGSKSVNMG-----YG 122
DB   91 -----gvkeqmfqlvgpvlencisgfussvayqsgsktytawgpangillechlg 144
QY   123 KEHGVIPTICDMFRINE-----LQDKNTCTCVBSVLEIYNNEVRDLLN- 169
DB   145 dqrgltpvrverifarikevmvgfvvlltlkssmlkgm-sitnvghysrfsdwlhy 203
QY   170 -----PSTKGML-----KVREHPSGPVEDLAKLVRSFOE 201
DB   204 lsiddlgrannrptgbekpgdkghvhrfslmchqdiredkvgyvenlteeyvkntid 263
QY   202 IENIMDEGNKARTVAATNNNETSSRSHAVFTLTQTOKWHDSETKMDTEKVKISLVDLAG 261
DB   264 vsqliklqngirtgatavntessrhcvftcvesrcknvadglssfktsrlnlvdlag 323
QY   262 SERATSGATGARLKEGAIEINRLSLTGRTVTAALADMS-SGKKQLVPYRDSVLTWLL 320
DB   324 serqtadaagaelkeagnlnarslgnlinllaieistqtkprh---ipydrsltfll 380
QY   321 KSLGSNSMTAMIAAISPADINFEEFLSTLYADSAAK 357
DB   381 qeslgsnaklamvcavspsqrs--ettatlrfaqrak 415

RESULT 15
GX   G40077
XX   G40077 standard; Protein; 1460 AA.
GX   G40077;
DX   18-OCT-2000 (first entry)
DX   Arabidopsis thaliana protein fragment SEQ ID NO: 49678.
KY   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
OS   Arabidopsis thaliana.
PN   EP1033405-A2.
PD   06-SEP-2000.
PX   25-FEB-2000; 2000EP-0301439.
XX   25-FEB-1999; 99US-0121825.
XX   05-MAR-1999; 99US-0123180.
PR   09-MAR-1999; 99US-0123548.
PR   23-MAR-1999; 99US-0125788.
PR   25-MAR-1999; 99US-0126264.
PR   29-MAR-1999; 99US-0126785.
PR   01-APR-1999; 99US-0127462.
PR   06-APR-1999; 99US-0128234.
PR   08-APR-1999; 99US-0128714.

```

PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144637.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145096.  
PR 22-JUL-1999; 99US-0145132.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0150520.  
PR 25-AUG-1999; 99US-0150520.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 28-SEP-1999; 99US-0155488.  
PR 28-SEP-1999; 99US-0155596.  
PR 29-SEP-1999; 99US-0155717.  
PR 04-OCT-1999; 99US-0157753.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160747.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160769.  
PR 21-OCT-1999; 99US-0160810.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match  
Rat Loc Similarity 27.98; Score 511.5; DB 21; Length 1460;  
Matches 130; Conservative 35.18; Pred No. 3.6e+01;  
Mismatches 60; Indels 73; Gaps 10;

QY 40 PPGAEKARKSGKTMDGKAFADPSYWSFKNAPYARQEDFDGLGVPLLDNAFG 99  
DB 9 ppsnplk-r-ksaetates-----gfds-----gvkeqmgfvgplvencslg 53  
QY 100 YNCCIFAYGQTGSGSKYSNMG-----YKHEGVIPRICQDMFRINE----- 141  
DB 54 fnsafyagtgsgskytmgwpgangllieehcggdgrgltprvferifarikemvpgivqf 113  
QY 142 ---LQDKNMLPCTFEVSYLEIYNERVOLLN-----PSTGNL 176  
DB 114 vvitlksnmlkgn-sitnvgahysrfsdwihylsldlqranrptgpkpogkhv 172  
QY 177 -----KYREHSTGYPVEDLAKLVRSFOEINLMDGCKARTVAATNNETSSRH 228  
DB 173 hrfsmchqrdedkgvymtceyykntdveqlilkgirntgatsvntesrah 232  
QY 229 AVFTLTQTKWDEETKMDTEKAKISLVDLAGSRATSTGATGARLKEAGETNSLSLT 288  
DB 233 cvtvcvsearcknvadgiisfksrslnlvdagserqkstgaagerlkeagnlnrsisl 292  
QY 289 GRVIAALADMS-SGCKKNOLPPYRDSVLTMLKSLGGSMTAMIAALSPADINFEETL 347  
DB 293 gnlnlrlaelstgkprh---lpyrderitflqeslgnaklamvcavspqrs--ettf 347  
QY 348 STLRYADSAK 357  
DB 348 stlrfaqrak 357

Search completed: April 25, 2001, 10:12:47  
Job time: 343 sec







D6 369 VSPSALNYEETI,STI.RYASRAR 390

```

1  RESULT          3
2  US-08-282-845-2
3  ; Sequences: 1, Application US/08/282845
4  ; Request No. 57191563
5  ; 1
6  ; GENERAL INFORMATION:
7  ;
8  ; APPLICANT: Reed, Steven G.
9  ; TITLE OF INVENTION: A 220kd Antigen Present in Leishmania
10 ;
11 ; TITLE OF INVENTION: Species
12 ;
13 ; NUMBER OF SEQUENCES: 3
14 ;
15 ; CORRESPONDENCE ADDRESS:
16 ;
17 ; ADDRESSEE: Immunex Corporation
18 ;
19 ; STREET: 51 University Street
20 ;
21 ; CITY: Seattle
22 ;
23 ; STATE: WA
24 ;
25 ; COUNTRY: USA
26 ;
27 ; ZIP: 98101
28 ;
29 ; COMPUTER READABLE FORM:
30 ;
31 ; MEDIUM TYPE: Floppy disk
32 ;
33 ; COMPUTER: Apple Macintosh
34 ;
35 ; OPERATING SYSTEM: Apple Macintosh Operating System 7.1a
36 ;
37 ; SOFTWARE: Microsoft Word for Macintosh 5.1a
38 ;
39 ; CURRENT APPLICATION DATA:
40 ;
41 ; APPLICATION NUMBER: US/08/282,845
42 ;
43 ; FILING DATE:
44 ;
45 ; CLASSIFICATION: 435
46 ;
47 ; PRIOR APPLICATION DATA:
48 ;
49 ; APPLICATION NUMBER: 08/006,676
50 ;
51 ; FILING DATE: JANUARY 15, 1993
52 ;
53 ; CLASSIFICATION: 435
54 ;
55 ; ATTORNEY/AGENT INFORMATION:
56 ;
57 ; NAME: Perkins, Patricia Anne
58 ;
59 ; REGISTRATION NUMBER: 34,693
60 ;
61 ; REFERENCE/DOCKET NUMBER: 5004-A
62 ;
63 ; TELECOMMUNICATION INFORMATION:
64 ;
65 ; TELEPHONE: (206)356-0430
66 ;
67 ; TELEFAX: (206)356-0644
68 ;
69 ; INFORMATION FOR SEQ ID NO. 2:
70 ;
71 ; SEQUENCE CHARACTERISTICS:
72 ;
73 ; LENGTH: 955 amino acids
74 ;
75 ; TYPE: amino acid
76 ;
77 ; TOPOLOGY: linear
78 ;
79 ; MOLSCULE TYPE: protein
80 ;
81 ; US-08-282-845-2

```

[illegible]

[illegible]

```

1  RESULT      5
2  PCT-US94-00324-1
3  : Sequence 1, Application PC/TUS9400324
4  : GENERAL INFORMATION:
5  : APPLICANT: Reed, Steven
6  : TITLE OF INVENTION: Diagnosis of Leishmaniasis
7  : NUMBER OF SEQUENCES: 3
8  : CORRESPONDENCE ADDRESS:
9  : ADDRESSEE: Immunex Corporation
10 : STREET: 51 University Street
11 : CITY: Seattle
12 : STATE: Washington
13 : COUNTRY: USA
14 : ZIP: 98101
15 : COMPUTER READABLE FORM:
16 : MEDIUM TYPE: floppy disk
17 : OPERATING SYSTEM: Apple Macintosh
18 : SOFTWARE: Microsoft Word, version 7.1
19 : CURRENT APPLICATION DATA:
20 : APPLICATION NUMBER: PCT/US94/00324
21 : FILING DATE:
22 : CLASSIFICATION:
23 : PRIOR APPLICATION DATA:
24 : APPLICATION NUMBER: US/08/006,676
25 : FILING DATE: 15-JAN-1993
26 : ATTORNEY/AGENT INFORMATION:
27 : NAME: Perkins, Patricia Anne
28 : REGISTRATION NUMBER: 34,693
29 : REFERENCE/DOCKET NUMBER: 5004-WO
30 : TELECOMMUNICATION INFORMATION:
31 : TELEPHONE: (206) 587-0430
32 : TELEFAX: (206) 587-0644
33 : INFORMATION FOR SEQ ID NO. 1:
34 : SEQUENCE CHARACTERISTICS:
35 : LENGTH: 955 amino acids
36 : TYPE: amino acid
37 : TOPOLOGY: linear
38 : MOLECULE TYPE: protein
39 : PCT-US94-00324-1

```

Query Match	38.5%;	Score 705.5;	DB 5;	Length 955;
Best Local Similarity	42.7%;	Pred. No. 6.9e-64;		
Matches 163: Conservative	58;	Mismatches 126;	Indels 35;	Gaps 8;

Qy	7	IKVYVYVBPINABE- -IDRACNCTVRMBCNQTIIIRP- - - - -PCAEKAKKSGXTIMD 57
Db	13	VKVSYVYVBPINABENNAPECTKYTVVAARQAAVTVVVLGSGNSGAAESKGTARVAQD 72
Qy	58	GPKAFADFSAYWSF- - -DKNAPNTAQEDFDLQGVLLPNDAKFGVNCNCFAYGOTGSK 114
Db	73	---FOEDHAWSVETPDCAGATATQADVFRTGCVPLVQHAQDFGNSCLFAYGOTGSK 128
Qy	115	SYSHMG- - - - -YKGEHVTIPRCDQKFRINELAKQKMLCTVYVYVYIENVRVDLL 168
Db	129	TYTHMGADVSLGSGENCVTPRCLBETFAKSAVPAQGHSHRWIVELVYGVYENVRVDLL 188
Qy	169	NPTSG- - - - -NLKVRBHPSTGYPYEDLAKLVINSPQVLEINLMDEGKATPVAATNM 220
Db	189	GKREKCYKGGEEVYVYREHSPYVYVLEEGORLYEVSQDLYVYVLELCMGVYRTASTKM 248
Qy	221	NETSSRSHAFATLTL- - - - -TKMDHPTKMDTKYKAKIISIDUAGSPRATSPGATGARLX 276

```

Db      249  NDRSSRSHAIIMLLREERTMTKSGTIRTAGKSSRNMLVDLAGSERVAGSQVEGQCFK 308
QY      277  EGAETNRSLSLTGRVIAALADMS-GRQKKNQLVPYRDSVLTWLLKDSIGNSMTAMIAA 335
Db      309  EATHINLSLTTLGRVIDVLADMATKGAKAQSVAPFRSKLTFILKDSIGNSKTFMIAT 368
QY      336  ISPADINFEETSLTRYADSAK 357
Db      369  VSPSALNVEETSLTRYASRAR 390

RESULT 6
US-08-713-815A-4
; Sequence 4, Application US/08713815A
; Patent No. 5830659
; GENERAL INFORMATION:
; APPLICANT: Russell J. Stewart
; TITLE OF INVENTION: ACTIVE MICROBUPLE-BASED
; TITLE OF INVENTION: SEPARATIONS BY KINESINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thorpe No. 5830659th & Western, L.L.P.
; STREET: 9035 South 700 East, Suite 200
; CITY: Sandy
; STATE: Utah
; COUNTRY: USA
; ZIP: 84070
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb storage
; COMPUTER: AST Ascentia 900N
; OPERATING SYSTEM: DOS 6.22
; SOFTWARE: Word Perfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,815A
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Alan J. Howarth
; REGISTRATION NUMBER: 36,553
; REFERENCE/DOCKET NUMBER: T3214/U-2202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (801)566633
; TELEFAX: (801)566 0750
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-713-815A-4

```

[illegible]

Db 280 KTH---IPYDSKLTRILOESLGNARTTIVICCSPASSENESEKSLTIDFGRAK 331

## RESULT 8

US-08-935-760-4

; Sequence 4, Application US/08935760A

; Patent No. 5952217

; GENERAL INFORMATION:

; APPLICANT: Corley, Neil C.

; APPLICANT: Corley, Neil C.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: Recombinant Yeast Cell and Assay Using Same

; FILE REFERENCE: ON0156sequence

; CURRENT APPLICATION NUMBER: US/08/935,760A

; CURRENT FILING DATE: 1997-09-23

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 545

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-935-760-4

Query Match 5.5%; Score 100; DB 2; Length 545;

Best Local Similarity 19.3%; Pred. No. 0.099;

Matches 77; Conservative 62; Mismatches 129; Indels 130; Gaps 18;

QY 6 NIKVVVRPNABDRGACIVRMENQTLTPPGAEKARKSG--KTIIDGPKAFA 63

Db 19 NSTWIGAGSKDAGTLANGSKPL-----PPNEEKKKDRPYSLIG----- 61

QY 64 FDRSVSFDKNAPYARQEDFDGLGVLPLDNAFKYNCCIFAYGOTGSKSYSMWYGYK 123

Db 62 -DKTNKKKEKREPLSPDFEHTIIV-----GFDVATGEF--TCMPEQWA----- 104

QY 124 EHGIVPRICQDMFRINELQDNKILACTIVEVSLEYIN-----ERVRLILN 169

Db 105 -----RLIQT--SNITSEQKNPQAVLDY--LEFYNSKTSNSQKYSFTDKSAEDYN 154

QY 170 PSTGKNL-K-VREHSTGPGYVED-----LAKLVRSFOEIMLME 208

Db 155 SSNALINRVANSETPAYPVYSEDEDDDDDDATPPPVITAPRPHTKSVYTHSV--IEPL--P 210

QY 209 GNKARTVAANTWNETSSRSRHVFTLTQKHDEETKMDTEKVAK--ISLVLAGSER-- 264

Db 211 VPTROVATSPISPTENTTPPDALTLTNEKKKKPKMSDEILEKLSRISVSGDEPKKY 270

QY 265 -----ATSTGATCARLKEGAENRSLSTLGRVTAALADMSSGCKKNOL----- 308

Db 271 TRFEKIGGACSGVTYVANDVATGQEV-----AIQMNLOQPKKELINEILV 318

QY 309 -----VPYDSVLV-----WLLKDSLGNMSTAMI 333

Db 319 MRENKNPNIVNYLDSYLVGDELWYVMEYLAGGSLTDVY 356

## RESULT 9

US-09-149-934-4

; Sequence 4, Application US/09149934B

; Patent No. 6139837

; GENERAL INFORMATION:

; APPLICANT: Corley, Neil C.

; APPLICANT: Corley, Neil C.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: ATP-DEPENDENT RNA HELICASE PROTEIN

; FILE REFERENCE: PE-0338-1 DIV

; CURRENT APPLICATION NUMBER: US/09/149,934B

; CURRENT FILING DATE: 1998-09-09

; EARLIER APPLICATION NUMBER: 08/892,256

; EARLIER FILING DATE: 1997-07-11

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PERL Program

; SEQ ID NO 4

; LENGTH: 746

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: GenBank ID No. G139837 1707046

US-09-149-934-4

Query Match 5.1%; Score 94; DB 4; Length 746;

Best Local Similarity 20.9%; Pred. No. 0.69;

Matches 76; Conservative 55; Mismatches 134; Indels 98; Gaps 18;

QY 18 AREIDRGAKIVYBMENQTLTPPGAEKARKSGKTIIDG--PK-----AFAFDRSYWS 70

Db 98 AKRKOKEASSAKSENATETTTKEPKCKRRKGGGENDGPKPKKREASKKTTDISANK 157

QY 71 FDKNAPYARQEDFDQDLGV-----PLLDNAPKYNCCIFAYGOTGSKSYSMWYGYK 122

Db 158 QFTFLPNEVL--EAIBQMGFSEPTIOSAVLPAAVRDRODVLGA-AETGCKT---LAFG 211

QY 123 KEHGVIPIRQ-----DMFRRINELQKKNLTC 150

Db 212 -----IPLVARLLESSDSQESTESTEVGRPRALIVAPTRVLVIQIMKHINALISTTOLIA 266

QY 151 TTVVSYTL-EIYREVRDILLAP-----STKGNL-KVREHSTGPGYV---EDLAKLVRSFOE 201

Db 267 TSVIGLQAQVKKERTVATN-KNETSSRSRSH--AVPTLTLTQKHDEETKMDTEKVAKIS-- 323

QY 202 TENLMDCKKARTVAATN-KNETSSRSRSH--AVPTLTLTQKHDEETKMDTEKVAKIS-- 255

Db 324 TDRWVEYFAELTHILNKIHSESEKELQTLVFSNATLTPAKAQDVAEEKKAKKELSSQ 383

QY 256 -----LYVDLAGSERATSTGATCARLKEG---AEINNSLSLSTLGRVTAALADMSSGCKKN 306

Db 384 OKTORILKLTG-----LRENHKRVLDLTROMGTAGCLVEARINCGNILLEKDT 430

QY 307 QLV 309

Db 431 SLV 433

RESULT 10

US-09-097-767A-26

; Sequence 26, Application US/09097767A

; Patent No. 6156232

; GENERAL INFORMATION:

; APPLICANT: Hey, Timothy D.

; APPLICANT: Owens Merlo, Ann

; APPLICANT: Walsh, Terence A.

; TITLE OF INVENTION: Nucleotide Sequences of Genes

; TITLE OF INVENTION: Encoding Sink Proteins and Uses Thereof for Improving the N

; TITLE OF INVENTION: Quality of Feeds

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dow AgroSciences, LLC

; STREET: 9330 Zionsville Road

; CITY: Indianapolis

; STATE: IN

; COUNTRY: US

; ZIP: 46268

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0,

; SOFTWARE: Version 1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/097,767A

; FILING DATE:

; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Borucki, Andrea T.  
 REGISTRATION NUMBER: 33651  
 REFERENCE/DOCKET NUMBER: 50451  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 317-337-4846  
 TELEFAX: 317-337-4847  
 INFORMATION FOR SEQ ID NO: 26:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 402 amino acids  
 TYPE: amino acids  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-097-767A-26

Query Match 5.0%; Score 92; DB 4; Length 402;  
 Best Local Similarity 22.4%; Pred. No. 0.41;  
 Matches 77; Conservative 43; Mismatches 127; Indels 96; Gaps 19;

QY 35 QVILTPPGAEE-----KARKSGKTIIDGPKAFADRSY-----WSFDKNAP 76  
 DB 60 QVLPPEKKVPELWFTEKTRSSITL-----AIRMNLYLVGFPTGGVWMEFGDG- 113  
 QY 77 NYARQEDLPFDGLVPLLDNAFKGYNCCIFAYGOTGSGKSYMMG-YGKEHGVIPIRCQDM 135  
 DB 114 -----DTHLLGDNPRLGFG-----GRYQDLGNKGLGVSMGR--AEM 150  
 QY 136 FRINELQKDKNLCTVEVSYLEIYNERNVRLDLSFSTGNLKVREHPTGPTVEDLAKLV 195  
 DB 151 THAVNDLAKKKAASITL-----KKVKRMQMPRAAKLAAAADPQADT-KSKLVKLV 202  
 QY 196 VRSFOEIMLMDENKARTVAANMMNNTSSKSHAVETLT-----QKNHDEETKMDTEK 250  
 DB 203 V-----WCVELRPNVST-YDAGFNSQBGV-TLTVTGKQVQKN-----DR 243  
 QY 251 VAKISLVLAGSERATSGTARGLKEAGAEINSLTGLRVIAALAMSSGQKKNLQVP 310  
 DB 244 ISKAQ-EWAHPVAPIDPMQKLGKDKNEARIVALKMOTTAAD-ASAKNKKKSM 301  
 QY 311 YDSYLTLLKDSGG-----NSMTAM-IAAISPADINFEE 345  
 DB 302 IGDDV-----PILGGINVGINSLVQDLARFAVQDHKKKE 338

RESULT 11  
 US-08-732-433-1  
 Sequence 1, Application US/08732433  
 Patent No. 6063611

GENERAL INFORMATION:  
 APPLICANT: van Solingen, Dieter  
 TITLE OF INVENTION: No. 6063611al Alkaline Cellulase and  
 TITLE OF INVENTION: Method of Producing Same  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genencor International, Inc.  
 STREET: 925 Page Mill Road  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304-1013  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/732.433  
 FILING DATE: 18-APR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US96/05651  
 FILING DATE: 26-APR-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Baumeister, Kirk  
 REGISTRATION NUMBER: 33,833

NAME: Stone, Christopher L.  
 REGISTRATION NUMBER: 35,696  
 REFERENCE/DOCKET NUMBER: GC287-US2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-846-7620  
 TELEFAX: 650-845-6504  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 574 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-732-433-1

Query Match 5.0%; Score 92; DB 3; Length 574;  
 Best Local Similarity 19.5%; Pred. No. 0.72;  
 Matches 66; Conservative 43; Mismatches 134; Indels 104; Gaps 13;

QY 28 IYRMEGNOITLTP-----PPGAEEKARKSGKTI--MDGPKAFADRSY--WSFDKNAP 76  
 DB 202 IVRSGSGONDRLPLVPTMETATSQLNNLYQTDIKLDDPNLIATVHYTGWPEFSVNA 261  
 QY 77 NYAR-QEDLPFDGLVPLLDNAFKGYNCCIFAYGOTGSGKSYMMG-YGKEHGVIPIR 130  
 DB 262 GYTRFEEDSKRE-----IETFDRVHHTFVARGIPVVLGFEGLGDFDKHTGVIQGEK 316  
 QY 131 IQCDMFRINELQKDKNLCTVEVSYL-----EYINERVRLDLSFSTGNLKV 177  
 DB 317 FEYLIHLNER-----DITHLMDNGQHFNRHITYEWDDEFLDLRASMGGRSS 366  
 QY 178 VRE-----HESTGPTVEDLAKLVRSPOEINLMDENKARTVAANMMNNTSSRSHAVETL 233  
 DB 367 VAESNFYTLKGDRADAT-----VTLQHGNETLGLQANGQLTPQDYTELNGE 416  
 QY 234 TLTKKHDEETKMDTEKVAKISLVLAGSERATSGTARGLKEAGAEINSLTGLRVIA 293  
 DB 417 RUTYKAH-----VLSIAGSGTLGTNGMVTAFENRGAADHFRVNT-----456  
 QY 294 ALADMSGQKKNLQVPYRDSVL-----TWLLKDSLGSNMTAMIA 334  
 DB 457 -----YRTEVLQSTQGHVSNFSPASFNGNSLATMEA 488

RESULT 12  
 US-08-368-281-2  
 Sequence 2, Application US/08368281  
 Patent No. 5721113  
 GENERAL INFORMATION:  
 APPLICANT: Libermann, Towia A  
 APPLICANT: Oetigen, Joerg P  
 APPLICANT: Kunsch, Charles A  
 TITLE OF INVENTION: NERF Genes  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SmithKline Beecham Corporation  
 STREET: 109 Swedeland Road  
 CITY: King of Prussia  
 STATE: Pennsylvania  
 COUNTRY: USA  
 ZIP: 19406  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/368.281  
 FILING DATE:  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Baumeister, Kirk  
 REGISTRATION NUMBER: 33,833

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5056
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-368-281-2

Query Match          5.0%; Score 91; DB 1; Length 521;
Best Local Similarity 19.9%; Pred. No. 0.78;
Matches 83; Conservative 61; Mismatches 145; Indels 128; Gaps 21;

QY 39 TPFGAEKARK-SGKTMDGPKAFADRSYWSF-----DKN-APNY-----ARQEDLFQ 86
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 SPELGIAKKAPREGKNTTY-----LWEFLDLQDKNTCPRIKWTQREKGFK 167
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 DLGVPLIDNAKGYNNCFAYGOTSGKSYSMG---YKKEGVIPRI-----QQD 134
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 168 -----LVDS--KAVSK-LAGKHKKPKDNYETMGALRYYYQGLLAKVGQRLVYQPKD 219
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 MFRRIEQLQDKNLTC-----TVEVSVLEIYNERVRLNPST----- 172
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 220 MPKNIWVDDDKSETCNEGLAGTTDEKSLERSLSAESLLKAASSVRSKNSPINCRA 279
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 -KGNLKV-----REHPSTGPVYVEDLAKLVRSFOEIEIENLMD-E-GNKARTVAAT 218
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 280 EKGVARVWITSPOCHDASSRPTTASVSATAAPRTVRVAMQPVVMTSLGQKISTVAVQ 339
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 219 NNN-----ETSSRSNAVFTLT-----TQKHDEETKMDTEKVAKISLVDLAGSE 263
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 340 SYNAGAPLITSTPTATSPKVIQTIPTVPWASTENGDKITMPAKIITIPATOLAQQ 459
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 RATESGATGA-----RLKEGAENR-----SLSLGRTVIALAD 297
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 400 LQTKSLNLTGSGSINVTGLAVRALTPVSIANGTPVRLSMPTQOASGQTPPVISAVI- 458
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 MSSGKKOKNOLVPRD--SVLTMLL---KDSLGGNSMTAMTALISPADINFEETLST 349
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 459 --KGAEVKSEAVAKQEHVDKTLLELVEEKPADGNKVTHTVVVVVSAPSALPVTMKT 513
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-08-368-281-4
; Sequence 4, Application US/08368281
; Patent No. 5721113
; GENERAL INFORMATION:
; APPLICANT: Libermann, Towia A
; APPLICANT: Oetting, Joerg P
; APPLICANT: Kunsch, Charles A
; APPLICANT: Winkler, Norf Genes
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08368,281
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Bruneister, Kirk
```

```

; REGISTRATION NUMBER: 33,833
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5056
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-368-281-4

Query Match          5.0%; Score 91; DB 1; Length 581;
Best Local Similarity 19.9%; Pred. No. 0.94;
Matches 83; Conservative 61; Mismatches 145; Indels 128; Gaps 21;

QY 39 TPFGAEKARK-SGKTMDGPKAFADRSYWSF-----DKN-APNY-----ARQEDLFQ 86
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 SPELGIAKKAPREGKNTTY-----LWEFLDLQDKNTCPRIKWTQREKGFK 227
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 87 DLGVPLIDNAKGYNNCFAYGOTSGKSYSMG---YKKEGVIPRI-----QQD 134
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 228 -----LVDS--KAVSK-LAGKHKKPKDNYETMGALRYYYQGLLAKVGQRLVYQPKD 279
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 MFRRIEQLQDKNLTC-----TVEVSVLEIYNERVRLNPST----- 172
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 280 MPKNIWVDDDKSETCNEGLAGTTDEKSLERSLSAESLLKAASSVRSKNSPINCRA 339
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 -KGNLKV-----REHPSTGPVYVEDLAKLVRSFOEIEIENLMD-E-GNKARTVAAT 218
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 340 EKGVARVWITSPOCHDASSRPTTASVSATAAPRTVRVAMQPVVMTSLGQKISTVAVQ 399
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 219 NNN-----ETSSRSNAVFTLT-----TQKHDEETKMDTEKVAKISLVDLAGSE 263
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 400 SYNAGAPLITSTPTATSPKVIQTIPTVPWASTENGDKITMPAKIITIPATOLAQQ 459
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 RATESGATGA-----RLKEGAENR-----SLSLGRTVIALAD 297
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 460 LQTKSLNLTGSGSINVTGLAVRALTPVSIANGTPVRLSMPTQOASGQTPPVISAVI- 518
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 MSSGKKOKNOLVPRD--SVLTMLL---KDSLGGNSMTAMTALISPADINFEETLST 349
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 519 --KGAEVKSEAVAKQEHVDKTLLELVEEKPADGNKVTHTVVVVVSAPSALPVTMKT 573
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-07-777-715-7
; Sequence 7, Application US/0777715
; Patent No. 5273889
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew
; APPLICANT: Campos, Manuel
; APPLICANT: Hughes, How P.A.
; TITLE OF INVENTION: CITOKINE-LEUKOTOXIN GENE FUSIONS AND
; METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07777,715
; FILING DATE: 19911016
; CLASSIFICATION: 424
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 29310-2001320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-777-715-7

Query Match 4.98; Score 90; DB 1; Length 1098;
Best Local Similarity 20.38; Pred. No. 3.4;
Matches 85; Conservative 70; Mismatches 167; Indels 96; Gaps 21;

Qy 7 IKVVVRVP---FNAREI---DRGAKIVRMGNOTILTPPPGAEKARKSGKTMDGP 59
Db 74 LEEVLNAPSKNLPREIKDSMDNKRITVLELQSGSETRFT---CEYDDATVNAVEFLN-- 128

Qy 60 KAFAPDRSWSFKNAPYARQEDLFODLGVPLLDNAFGYNNCIFA--YGQTSGSKSYS 117
Db 129 KWIIFQCSYIS-----TMTGDLSPRLTTLNGLKNTLTATKSLHKGQSULT 176

Qy 118 MMGYGKEGV-----IPRCQDMFRRINELQDKNLTCTVEVSYLEIYNERVRLDNFS 171
Db 177 QAGSSSLKTGAKKILLYIPONTDYDEOQNGLQ---DLVKAEEGLIEVQREERNIATAQ 233

Qy 172 TK-GNLKV-----REHSTGPYVEDLAKLV--RSFOELENLMDGKNKARTV----- 215
Db 234 TSLGTGTTAIGLTERGIVLSAPOIDKLLQKTAGQALGSAESIYONAKAKTVLSGIQSI 293

Qy 216 -----AATNMNE---TSSRSNAVFTLTITOKWHEETKMDTEKVA-KISLVDLAQS 305
Db 348 FGSKLQNIKGSLTGDKLNIGGLDKAGLDVIGSLGSGATAALVLADKNASTAKVGA 407

Qy 306 -----NQLV-PYRDSVLTWLLKQSL--GGSNWTAMTAISPADINPETLTSLRYADSA 356
Db 408 GFELANQVVGWNTKAVSSITLAORVAAGLSTGCPVAALIASTVSL--AISPLAFAGTA 463

RESULT 15
US-08-170-126-2
; Sequence 2, Application US/08170126
; Patent No. 5594107
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW
; APPLICANT: CAMPOS, MANUEL
; APPLICANT: HUGHES, HUW P. A.
; TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,126
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,715
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/571,301
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0013.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-170-126-2

Query Match 4.98; Score 90; DB 1; Length 1098;
Best Local Similarity 20.38; Pred. No. 3.4;
Matches 85; Conservative 70; Mismatches 167; Indels 96; Gaps 21;

Qy 7 IKVVVRVP---FNAREI---DRGAKIVRMGNOTILTPPPGAEKARKSGKTMDGP 59
Db 74 LEEVLNAPSKNLPREIKDSMDNKRITVLELQSGSETRFT---CEYDDATVNAVEFLN-- 128

Qy 60 KAFAPDRSWSFKNAPYARQEDLFODLGVPLLDNAFGYNNCIFA--YGQTSGSKSYS 117
Db 129 KWIIFQCSYIS-----TMTGDLSPRLTTLNGLKNTLTATKSLHKGQSULT 176

Qy 118 MMGYGKEGV-----IPRCQDMFRRINELQDKNLTCTVEVSYLEIYNERVRLDNFS 171
Db 177 QAGSSSLKTGAKKILLYIPONTDYDEOQNGLQ---DLVKAEEGLIEVQREERNIATAQ 233

Qy 172 TK-GNLKV-----REHSTGPYVEDLAKLV--RSFOELENLMDGKNKARTV----- 215
Db 234 TSLGTGTTAIGLTERGIVLSAPOIDKLLQKTAGQALGSAESIYONAKAKTVLSGIQSI 293

Qy 216 -----AATNMNE---TSSRSNAVFTLTITOKWHEETKMDTEKVA-KISLVDLAQS 305
Db 348 FGSKLQNIKGSLTGDKLNIGGLDKAGLDVIGSLGSGATAALVLADKNASTAKVGA 407

Qy 306 -----NQLV-PYRDSVLTWLLKQSL--GGSNWTAMTAISPADINPETLTSLRYADSA 356
Db 408 GFELANQVVGWNTKAVSSITLAORVAAGLSTGCPVAALIASTVSL--AISPLAFAGTA 463
```

Search completed: April 25, 2001, 10:13:24  
JOB Time: 240 sec





**THIS PAGE BLANK (USPTO)**

Query Match	62.5%	Score 1147	DB 2	Length 1695
Best Local Similarity	61.8%	RefSeq No. 6,9e-78		
Matches 225	Conservative 56	Mismatches 57	Indels 24	Gaps 6
QY	4	GNIKVAVVRPENNAREIDRGAQCTVYECNQTILTPPGAPAEKARKSGKTIIDGPKAPA	63	
DB	4	:       :       :       :       :       :       :       :       :		
QY	3	GASVYAVVRPENSREMSRSDSKLIQSSGTTIYNYPQPKET-----PKSPS	51	
DB	3	:       :       :       :       :       :       :       :       :		
QY	64	FDRSVYSEFKDAP--NYARQEDLPDGLVPLDPAKGYNNKIFAYGQTSGSKSYNMK	120	
DB	64	:       :       :       :       :       :       :       :       :		
QY	52	FDYSYWS--HTSPEDIINTASQKQVRDIEGEMLRHAFEGYVNCIFYGQTGACKSYNMK	109	
DB	52	:       :       :       :       :       :       :       :       :		
QY	121	YK--EHGVIPTICDMPFRINELQKDKNLITTVESYLETYNERNVEDLLMPSTKGNLKV	178	
DB	121	:       :       :       :       :       :       :       :       :		
QY	110	KQEKDOOGIIFQCELFDRINDTND--NMSYSVSVSYNICYTCEVRDLNPKNKGRLV	168	
DB	110	:       :       :       :       :       :       :       :       :		
QY	179	REHPSTGPVIEDLAKLVFASFOE:ENLMDEGNKARTVAATNNHSTSSRSRAVFTLTQK	238	
DB	179	:       :       :       :       :       :       :       :       :		
QY	169	REHPLDGPVIEDLSKLAVSYND:LOQLMDGKNPRTVAATNNHSTSSRSRAVFIIFTQK	228	
DB	169	:       :       :       :       :       :       :       :       :		
QY	239	WHDFTKMDTKVAKISLVDLAGSERATSGATGARLKEGATNRSJLSTGRVIAALADM	298	
DB	239	:       :       :       :       :       :       :       :       :		
QY	229	RHDAETITTEKYSKISLVDLAGSERADSTGAGRLKEGATNRSJLSTGRVIAALAEK	288	
DB	229	:       :       :       :       :       :       :       :       :		
QY	299	SSG-----KQKNQLVYPRDSVTWLLDKSLGGNSNTANIAALSPADINFEETLSTLYRA	353	
DB	299	:       :       :       :       :       :       :       :       :		
QY	289	DGSPNKKKKKKTDIFPYDSVTWLLREALGGNRTNVAALSPADINFEETLSTLYRA	348	
DB	289	:       :       :       :       :       :       :       :       :		
QY	354	DSAK	357	
DB	354			
QY	349	DRAK	352	
DB	349			

A:introns: 34/1; 84/2; 140/3; 200/2; 285/3; 339/2; 490/2; 580/1; 631/3; 672/1; 738/2;  
C:Superfamily: kinesin-related protein unc-104; kinesin motor domain homology; plecks

RESULT 2  
JN0114  
Kinesin-related protein unc-104 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 29-May-1998  
C:Accession: JN0114  
R:Otsuka, A.J.; Jeyaprakash, A.; Garcia-Anoveros, J.; Tang, L.Z.; Fisk, G.; Hartshorne, Neuron 6, 113-122, 1991  
A:Title: The C. elegans unc-104 gene encodes a putative kinesin heavy chain-like protein  
A:Reference number: JN0114; MUID:91097805  
A:Accession: JN0114  
A:Molecule type: mRNA  
A:Residues: 1-1584 <OTS>  
A:Cross-references: GB:M58582  
A:Note: 598-Thr and 930-Met were also found  
C:Genetics:  
A:Gene: unc-104  
C:Superfamily: kinesin-related protein unc-104; kinesin motor domain homology; pleckstrin  
C:Keywords: ATP; microtubule binding; P-loop  
F:4-353/Domain: kinesin motor domain homology <NMOT>  
F:93-100/Region: nucleotide-binding motif A (P-loop)  
F:1285-1287/Region: cell attachment (R-G-D) motif  
F:99/Binding site: ATP (Lys) #status predicted

Query Match 60.3%; Score 1106.5; DB 1; Length 1584;  
Best Local Similarity 61.2%; Pred. No. 7e-75;  
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

QY 6 NIKVVRVPPNAREIDRGAKCIVRMGNGQITLTPPGAEEKARKSGKTMGPKAFAD 65  
DB 3 SVKAVVRPPNOREISNTSKVLQVNGNTT-----NGHSINKNFSEFND 50  
QY 66 RSYWFDKNAPNARQEDLDQGLVPLDPAFGYNNCFAYGQTGSGKSYMMGYCK-- 123  
DB 51 HSYWSEARNDFHPTQKVYELGVEMLEHAFEGYNVCIFYGQTGSGKSYMMGKANOP 110  
QY 124 -EHGVIPRICODMFRINELQKKNLTCTVEVSYLEIYNERNVDLLNPSTKGNLKVREHP 182  
DB 111 DMGIIPRLCNDFARIDN--NNQDKVQSVSEVSYMEIYCERVKDLLNPNSGGNLRVREHP 169  
QY 183 STGYPVEDLAKLVRSFOEINLMDGKNGKARTVAATNNNETSSRSRSHAVFTLTQKHDE 242  
DB 170 LLGYPVDDLTVMKVCYHDICNLMDGKNGKARTVAATNNNETSSRSRSHAVFTLTQKHCA 229  
QY 243 ETQMDETKAKISLVLAGSRATSTGATGARLKEGAINRSLTSLGRVTAALADMSGK 302  
DB 230 DSNIDTEKHSKISLVLAGSRANSTGAQRLEKGANINKSLTTLGLVSKLAESTRK 289  
QY 303 OKKNO-LVPIYRDSVLTLLRENLGGNSMTAMIAASPADINFEETLTSLRYADSAK 357  
DB 290 KKSNGKVIPIYRDSVLTLLRENLGGNSKTMALAASPADINFDETLTSLRYADRAK 345

RESULT 3  
T15822  
Kinesin-like protein unc-104 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-Jan-2000  
C:Accession: T15822  
R:Du, Z.  
submitted to the EMBL data library, February 1996

A:Description: The sequence of C. elegans cosmid C52E12.  
A:Reference number: 218412  
A:Accession: T15822  
A:Status: preliminary;  
A:Molecule type: DNA  
A:Residues: 1-1584 <DUZ>  
A:Cross-references: EMBL:U50135; NID:g1208884; PID:g1208885; PID:AAA93453.1; CESP:unc-104  
C:Genetics:  
A:Gene: CESP:unc-104

Query Match 60.3%; Score 1106.5; DB 2; Length 1584;  
Best Local Similarity 61.2%; Pred. No. 7e-75;  
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

QY 6 NIKVVRVPPNAREIDRGAKCIVRMGNGQITLTPPGAEEKARKSGKTMGPKAFAD 65  
DB 3 SVKAVVRPPNOREISNTSKVLQVNGNTT-----NGHSINKNFSEFND 50  
QY 66 RSYWFDKNAPNARQEDLDQGLVPLDPAFGYNNCFAYGQTGSGKSYMMGYCK-- 123  
DB 51 HSYWSEARNDFHPTQKVYELGVEMLEHAFEGYNVCIFYGQTGSGKSYMMGKANOP 110  
QY 124 -EHGVIPRICODMFRINELQKKNLTCTVEVSYLEIYNERNVDLLNPSTKGNLKVREHP 182  
DB 111 DMGIIPRLCNDFARIDN--NNQDKVQSVSEVSYMEIYCERVKDLLNPNSGGNLRVREHP 169  
QY 183 STGYPVEDLAKLVRSFOEINLMDGKNGKARTVAATNNNETSSRSRSHAVFTLTQKHDE 242  
DB 170 LLGYPVDDLTVMKVCYHDICNLMDGKNGKARTVAATNNNETSSRSRSHAVFTLTQKHCA 229  
QY 243 ETQMDETKAKISLVLAGSRATSTGATGARLKEGAINRSLTSLGRVTAALADMSGK 302  
DB 230 DSNIDTEKHSKISLVLAGSRANSTGAQRLEKGANINKSLTTLGLVSKLAESTRK 289  
QY 303 OKKNO-LVPIYRDSVLTLLRENLGGNSMTAMIAASPADINFEETLTSLRYADSAK 357  
DB 290 KKSNGKVIPIYRDSVLTLLRENLGGNSKTMALAASPADINFDETLTSLRYADRAK 345

RESULT 4  
A55289  
Kinesin-like protein KIF1B - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: A55289  
R:Nangaku, M.; Sato-Yoshitake, R.; Okada, Y.; Noda, Y.; Takemura, R.; Yamazaki, H.; H  
Cell 79, 1209-1220, 1994  
A:Title: KIF1B, a novel microtubule plus end-directed monomeric motor protein for tra  
A:Reference number: A55289; MUID:95094296  
A:Accession: A55289  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1150 <NAN>  
A:Cross-references: GB:D17577; NID:g407338; PIDN:BAA04503.1; PID:g407339  
C:Superfamily: kinesin-related protein KIF1B; kinesin motor domain homology  
C:Keywords: P-loop  
F:6-354/Domain: kinesin motor domain homology <NMOT>  
F:97-104/Region: nucleotide-binding motif A (P-loop)

Query Match 60.3%; Score 1106; DB 1; Length 1150;  
Best Local Similarity 61.1%; Pred. No. 4.8e-75;  
Matches 218; Conservative 59; Mismatches 64; Indels 16; Gaps 5;

QY 4 GNTKVVVRPPNAREIDRGAKCIVRMGNGQITLTPPGAEEKARKSGKTMGPKAF 63  
DB 3 GASKVAVVRPPNRSRESKSCIIQMGNGSTSIINKNPKE-----APKSFS 51  
QY 64 FDRSYWSP--DNKAPNARQEDLDQGLVPLDPAFGYNNCFAYGQTGSGKSYMMGYG 122  
DB 52 FDSYSHTSFSDPCFQSONRVNIDKEMLLHAFEGYNVCIFYGQTGSGKSYMMGKQ 111  
QY 123 KEHG--VPIRICODMFRINELQKKNLTCTVEVSYLEIYNERNVDLLNPSTKGNLKVRE 180  
DB 112 BESQAVIIPQICELETFEKIND--NCNEWSYSEVSYMEIYCERVKDLLNPNSGGNLRVRE 170  
QY 181 HSTGYPVEDLAKLVRSFOEINLMDGKNGKARTVAATNNNETSSRSRSHAVFTLTQKH 240  
DB 171 HPLGLPYVEDLSKLVATSYVDIADLMDAGNGKARTVAATNNNETSSRSRSHAVFTLTQKH 230

		QY	241	DEFTWMTOKTEKAKISILVDLAGSERSTGATCARLKEGAFFINBSLSITGLVRYAALADWS	300
		Db	231	DPETNLSTEKYKSILVDLAGSERADSTGAGTKRLEKANINSLTTLCKVISALAEYSK	290
		QY	301	GOKRKNOVLPRVSIVTLLWKLOJSGNSKMTAAIASPADINFEETLSTRYADSAK	357
		Db	291	KKKKTDFIPYRSVSLVTLLRENLGGNSRTMAVAASPADINTDETLSLTRYADRAK	346
		RESULT	5		
		T31827			
		kinesin-73 - fruit fly ( <i>Drosophila melanogaster</i> )			
		C:Species: Drosophila melanogaster			
		C>Date: 20-Sep-1999			
		C:Accession: T13827			
		R:L.I., H.P.; Liu, Z.M.; Nirenberg, M.			
		Proc. Natl. Acad. Sci. U.S.A. 94, 1086-1091, 1997			
		A>Title: Kinesin-73 in the nervous system of <i>Drosophila</i> embryos .			
		A:Reference number: 217784; PMID:97188425			
		A:Accession: T13827			
		A>Status: preliminary.			
		A:Translation: translated from GB/EMBL/DDBJ			
		A:Residues: 1-1921 <LINK>			
		A:Cross-references: EMBL:U81788; NID:g1906595; PID:g1906596; PIDD:AAB50404.1			
		C:Genetics:			
		A/Cross-references: FlyBase:Fgn0019968			

Query Match	56 18:	Score 1028:	DB 2:	Length 1921:
Best Local Similarity	60 35:	Pred. Neg 7, 4C-65:		
Matches 213:	Conservative 41:	Mismatches 85:	Indels 10:	Gaps 4:
QY	7	IKVVYVRFPNAREIDRGAKCAGTIVEMEGNOTITLTPPCABEKARKSGKTIWQPKAFADR	56	
DB	6	IKVAVVRFPNREIELDTKCIIVMEKQOITLQNPPLLEKTERQ-----PKTAFDTH	58	
QY	67	SYWSDKNAPYARQEDLPDGLVPLNDLAFQYNNCTIPAYGTGSGKSTSMNCTGCKBERG	126	
DB	59	CPTSLNPFDEPNASQETVFDYGVGRGLDNAAFQYGNACIPAYGTGSGKSTYTMGTQESKG	118	
QY	127	VIPRICDMFRRLNQLQKDKNLTCTVEYSVLEITNVERVDLNPSS--TKGNLKVRHPSGTG	185	
DB	119	IPIPRCQDLFASIAN--KSTPELTKYVEYSVMEYNEKVHLDLPKPKOSLKVRHYNMG	177	
QY	186	YPYEDLAKLYVRSGEITENLADGCKARTVAATNNETSSSHSAVFTLTLTQKHWDSTK	245	
DB	178	YPVQGLSAVLTQSDINLTALGSGARKLKECATINSELSTLGRVYATLADNDSQKOKK	305	
QY	246	MDTPEKVAATLSVLQVADSTGATGCAKARKEATINSELSTLGRVYATLADNDSQKOKK	305	
DB	238	VSGENKVSBNLSVLQAGSERAVTICAVGDGRLKSGINSKSLTTLTGILSKLQDSNGKSG	297	
QY	306	N-OLVPPVRSYLTWLKDSLGNSMTAAIISPADINPEPTLSTLTVADAAK	357	
DB	298	NDKFPVPPVRSYLTWLKONLGNSKRTYVWATISFSAADNYETPTLSTLTVADAAK	350	

RESULT 6  
T29237  
hypothetical protein F56E3.3 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T29237  
R:Gattung, S.; Wu, X.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of *C. elegans* cosmid F56E3.  
A:Reference number: Z20592  
A:Accession: T29237  
A:Status: preliminary; translated from GI/F56E3/DD83  
A:Molecule type: DNA  
A:Residues: 1-1576 <GAT>  
A:Cross-references: EMBL:U41536; PIDB:AA852613.1; GSPOB:GN00028; CESP:F56E3.3

A: Experimental source: strain Bristol N2; clone F56E3  
C: Genetics:  
A: Gene: CESP.F56E3.3  
A: Map position: X  
A: Introns: 23/1; 47/2; 125/3; 162/2; 191/3; 214/2; 295/3; 498/3; 532/1; 651/3; 723/2;

Query Match 52.6%; Score 965.5; DB 2; Length 1576;  
Best Local Similarity 53.9%; Pred. No. 2.8e-64;  
Matches 208; Conservative 45; Mismatches 82; Indels 51; Gaps 6;

66	QY	7	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	66
67	QY	8	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	67
68	QY	9	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	68
69	QY	10	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	69
70	QY	11	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	70
71	QY	12	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	71
72	QY	13	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	72
73	QY	14	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	73
74	QY	15	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	74
75	QY	16	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	75
76	QY	17	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	76
77	QY	18	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	77
78	QY	19	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	78
79	QY	20	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	79
80	QY	21	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	80
81	QY	22	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	81
82	QY	23	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	82
83	QY	24	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	83
84	QY	25	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	84
85	QY	26	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	85
86	QY	27	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	86
87	QY	28	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	87
88	QY	29	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	88
89	QY	30	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	89
90	QY	31	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	90
91	QY	32	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	91
92	QY	33	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	92
93	QY	34	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	93
94	QY	35	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	94
95	QY	36	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	95
96	QY	37	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	96
97	QY	38	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	97
98	QY	39	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	98
99	QY	40	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	99
100	QY	41	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	100
101	QY	42	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	101
102	QY	43	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	102
103	QY	44	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	103
104	QY	45	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	104
105	QY	46	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	105
106	QY	47	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	106
107	QY	48	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	107
108	QY	49	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	108
109	QY	50	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	109
110	QY	51	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	110
111	QY	52	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	111
112	QY	53	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	112
113	QY	54	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	113
114	QY	55	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	114
115	QY	56	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	115
116	QY	57	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	116
117	QY	58	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	117
118	QY	59	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	118
119	QY	60	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	119
120	QY	61	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	120
121	QY	62	IKVYVVRVPPNAREIDRACATVEMGNOTLTPPPCAEAKRKSQITMDGPKAFADP	121
122	QY			

RESULT 7

Ti6759

hypothetical protein R144.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: Ti6759

R:Favelllo, T.

submitted to the EMBL Data Library, March 1995

A:Description: The sequence of C. elegans cosmid R144.

A:Reference number: 218571

A:Accession: Ti6759

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-328 <FAV>

A:Cross-references: EMBL:U23515; NID:G746493; PID:G746493; CESP:R144.1

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:R144.1

A:Introns: 80/2: 125/1: 261/2: 331/3: 350/2: 516/3: 589/1: 615/2: 638/2: 726/1: 865/3

Query Match	47.9%	Score 878.5;	DB 2;	Length 928;
Best Local Similarity	51.1%	Pred. No. 4.6e-58;		
Matches 189;	Conservative 54;	Mismatches 100;	Indels 27;	Gaps 6;

[illegible]

QY 110 TSGSKSYSMKGYKEHGVIPRICODMFRINELQKOKNLCTCTVEYSVLEIYNERYDRLN 169  
 DB 109 TSGSKSYSMKGYKEHGVIPRICODMFRINELQKOKNLCTCTVEYSVLEIYNERYDRLN 167  
 QY 170 --PSTGNLKYVREBPSTGPYVEDLAKLVVRSFOENIMDEGNKARFVAATWAFETSSRS 227  
 DB 168 STTPPGGLKLYREHPKNGFYENLTVPVSPFEIANKIEGTSKRIINATQMNATSSRA 227  
 QY 228 HAVFTLTQKWHDEETKMDTEKVAKISLVDLGASERATSTGATCARLKEGATNRSLSIT 287  
 DB 228 HTIVKTFNQK--SSKAGGTSMKKSEINLVDLGASERATSTGATCARLKEGATNRSLSIT 286  
 QY 288 LGRVIALADMSSCKOKKKNOLVPRDVSYVLLWLLKSLGNSMTAMIAISPADINFEETL 347  
 DB 287 LGRVIALHDSQAKSGKKTQIPYRDSVLTCLKNALGNSKTIIMIAISPADINFEETL 346  
 QY 348 STLRYADSAK 357  
 DB 347 STLRFADRAK 356

RESULT 8  
 T13750  
 Kinesin-like protein 38B - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*  
 C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C:Accession: T13750  
 R:Ohkura, H.; Torok, T.; Tick, G.; Hohnsels, J.; Kiss, I.; Glover, D.M.  
 J. Cell Sci. 110, 945-954, 1997  
 A:Title: Mutation of a gene for *Drosophila* kinesin like protein, KLP38B, leads to fail  
 A:Reference number: T17752; MUID:97296456  
 A:Accession: T13750  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1121 <OK>  
 A:Cross-references: EMBL:Y10667; NID:g2253151; PIDN:CAA71675.1; PID:g2253152  
 C:Cross-references: FlyBase:FBgn0004374  
 A:Map position: 2  
 C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match 43.3%; Score 793.5; DB 2; Length 1121;  
 Best Local Similarity 44.4%; Pred. No. 1.5e-51;  
 Matches 181; Conservative 56; Mismatches 114; Indels 57; Gaps 9;  
 QY 1 MGGGNKVVVRVPFNAEIDRG-AKCIVRMEGNQITLTPPPGAEBKARKSGKTMDGP 59  
 DB 115 VSESNMIVVRVPLNALECTRGQVTNVVGHGNSNELTVQAGSSADA-SAGVT----- 168  
 QY 60 KAPAFDSWFDKNAPYARQEDLQDGLVPLLDNAFKYNNCTIFAYGTGSGKSYSM 119  
 DB 169 HFSIYVYSCDPERKNFAQAKVEGTARPLIDTAFEGYNACLFAYGTGSGKSYSM 228  
 QY 120 GY-----GKEH--GVPRICDMFRNELQDKNLCTCTVEYSVLEIYNERYD 166  
 DB 229 GLEALDMLGPPHDEAGIIPFCHELFRRIEAVKSOQLOVEYSVFEIYNEKHID 288  
 QY 167 LNL-----PSTGNLKYVREBPSTGPYVEDLAKLVVRSFOENIMDE 208  
 DB 289 LLSVQHAARATGESTPIQOOQOORPALKVREHPFGYVVDLSAHSVDSISALRNMLAV 348  
 QY 209 GNKARTVAATNNETSSSHAVF--TLTLTKQKHDEETKMDTE-----KVAKISL 256  
 DB 349 GNSQAFSTAMNDKSSRSHTFNVLNLTDLSSDGLSDTSSASSLRQTRRSKLSI 408  
 QY 257 VDLGASERATSTGATCARLKEGATNRSITGLGVYIALADMS-----GKQKNOLY 309  
 DB 409 VDLGASERISVSGSNGERIREGYSINKSLTLGKYVIALADSKASPLSGTPTFTY 468  
 QY 310 PYRDSVLTWLLKSLGNSMTAMIAISPADINFEETLSTLRYADSAK 357  
 DB 469 PYRDSVLTWLLRNLGNSKTVMLATISPASIHADETLATRYACKAR 516

RESULT 9  
 T13796  
 Kinesin-related protein 38B - fruit fly (*Drosophila melanogaster*) (fragment)  
 C:Species: *Drosophila melanogaster*  
 C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
 C:Accession: T13796  
 R:Molina, Y.; Baars, S.; Hales, K.; Fuller, M.T.; Ripoll, P.  
 J. Cell Biol. 139, 1361-1371, 1997  
 A:Title: A chromatin associated kinesin-related protein required for normal mitotic  
 A:Reference number: Z17759; MUID:98060833  
 A:Accession: T13796  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1121 <MOL>  
 A:Cross-references: EMBL:Y15247; NID:g2578010; PIDN:CAA75531.1; PID:g2578011  
 C:Cross-references: FlyBase:FBgn0004374  
 A:Gene: tlo  
 C:Superfamily: kinesin-related protein Eg5; kinesin motor domain homology

Query Match 43.2%; Score 792.5; DB 2; Length 1121;  
 Best Local Similarity 44.7%; Pred. No. 1.8e-51;  
 Matches 180; Conservative 55; Mismatches 111; Indels 57; Gaps 9;  
 QY 6 NIKVVVRVPFNAEIDRG-AKCIVRMEGNQITLTPPPGAEBKARKSGKTMDGPKAPAF 64  
 DB 120 NMIVAVRVRPLNALECTRGQVTNVVGHGNSNELTVQAGSSADA-SAGVT-----HFFSY 173  
 QY 65 DRSYWSFDKNAPYARQEDLQDGLVPLLDNAFKYNNCTIFAYGTGSGKSYSMGY--- 121  
 DB 174 DOVYISCDPERKNFAQAKVEGTARPLIDTAFEGYNACLFAYGTGSGKSYSMGIEAL 233  
 QY 122 -----GKEH--GVPRICDMFRNELQDKNLCTCTVEYSVLEIYNERYDRLN-- 169  
 DB 234 DDAALDGGPPHYEAGIIPFCHELFRRIEAVKSOQLOVEYSVFEIYNEKHIDLSVQ 293  
 QY 170 -----PSTGNLKYVREBPSTGPYVEDLAKLVVRSFOENIMDEGNKAR 213  
 DB 294 HAAATGESTPIQOOQOORPALKVREHPFGYVVDLSAHSVDSISALRNMLAVGNSOR 353  
 QY 214 TVAATNNETSSSHAVF--TLTLTKQKHDEETKMDTE-----KVAKISLVLAG 261  
 DB 354 ATATMNDKSSRSHTFNVLNLTDLSSDGLSDTSSASSLRQTRRSKLSLVLAG 413  
 QY 262 SERATSTGATCARLKEGATNRSITGLGVYIALADMS-----GKQKNOLPYRDS 314  
 DB 414 SERISVSGSNGERIREGYSINKSLTLGKYVIALADSKASPLSGTPTFTYRDS 473  
 QY 315 VLTWLLKSLGNSMTAMIAISPADINFEETLSTLRYADSAK 357  
 DB 474 VLTWLLRNLGNSKTVMLATISPASIHADETLATRYACKAR 516

RESULT 10  
 S38982  
 Kinesin-related protein KRP85 - sea urchin (*Strongylocentrotus purpuratus*)  
 N:Alternate names: kinesin-2 chain A; KRP (85/95) 85K chain  
 C:Species: *Strongylocentrotus purpuratus* (purple urchin)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: S38982; S72551  
 C:Colo: J.G. Thompson, 1993; Wedeman, K.P.; Hall, K.; Vuong, T.; Scholey, J.M.  
 Nature 366, 268-270, 1993  
 A:Title: Novel heterotrimeric kinesin-related protein purified from sea urchin eggs.  
 A:Reference number: S38982; S38982; MUID:94050179  
 A:Accession: S38982  
 A:Molecule type: mRNA  
 A:Residues: 1-1121 <COLL>  
 A:Cross-references: EMBL:L16993; NID:g295245; PIDN:AAA16098.1; PID:g295246  
 A:Accession: S72551  
 A:Molecule type: Protein

A: Residues: 2-5,'X','7-11,59-64,125-132,222-226,'X',228-230 <COL2>  
 C: Complex: heterotrimer of a 115K chain and two kinesin-related chains of 95K (PIR:S5869)  
 C: Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
 C: Keywords: ATP; heterotrimer; microtubule binding; P-loop  
 F:11-348/Domain: kinesin motor domain homology <KMT>  
 F:97-104/Region: nucleotide-binding motif A (P-loop)  
 F:103/Binding site: ATP (Lys) #status predicted

Query Match 40.5%; Score 743.5; DB 1; Length 699;  
 Best Local Similarity 46.0%; Pred. No. 14e-48  
 Matches 167; Conservative 62; Mismatches 99; Indels 35; Gaps 10;

QY 2 SGGNTKVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 1 KVVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 6 SGNDVRVVVRCRPLNSKETSGFKSVYKMDMRGTGVQVTPNA-----PSG-----EPPK 56  
 QY 61 AFAEDTSYSGFNAPYARQEDFDGLGVPLLDNAFGYNNCFAYGOTSGSKYSNMG 120  
 Db 1 KVVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 57 SFTPTVTF-----APG-AKOTDYNVNOTPAPVDVIAIEGYNGTIFAYGOTGKTTFMEG 109  
 QY 121 YGKE---HGVIPRICQDMFRINELQKDKMLCTVEYSLEYIYNERVRLNFTSKGNLK 177  
 Db 1 KVVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 110 VRSOPELRGIIIPNSFAHIFGH--AKQGVNVLRYVRSLEYIYNEVKOLLGKDDQOHRLE 167  
 QY 178 VREHPSTGPVVEDLAKLVRSFOETENLMDGNKARTVAATNNNNTSSSRSHAVTLTITQ 237  
 Db 1 KVVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 168 VKEPQVGVYKQVLSAAYNADMDRIITLGNKMSGATNNNESSSRSHAFITTLIER 227  
 QY 238 KWHDEETKQTE---KVKISIVDLASGRATSGTGAIRKEGAENISLSTGLGWTA 294  
 Db 1 KVVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 228 ----SDWLDKEQVRYVGLKHWDLASGERQTKGTGAORLKEATKINLSLSTGLNVISS 283  
 QY 295 LADMSSGQKKQLPVPRDYSVLTMLKDSLGNSMTAMTAISPADINFEETLSTLYRAD 354  
 Db 1 KVVVVVRPPNAREIDGAKCIVRM-EGSNOTLTTPPGAEEKARKSGKTMGPK 60  
 Db 284 LVD-----GKSTHPIYRNSKTLRLQDSLGNAKTVMCANIGPAENYDETISTLYRAN 337  
 QY 355 SAK 357  
 Db 338 RAK 340

RESULT 11  
 S58691  
 Kinesin-related protein KRP95 - sea urchin (Strongylocentrotus droebachiensis)  
 A: Residues: 1-742 <RAS>  
 C: Superfamily: kinesin-related protein KIF3; kinesin motor domain homology  
 C: Keywords: ATP; heterotrimer; microtubule binding; P-loop  
 F:9-345/Domain: kinesin motor domain homology <KMT>  
 F:95-102/Region: nucleotide-binding motif A (P-loop)  
 F:103/Binding site: ATP (Lys) #status predicted

Query Match 38.8%; Score 712; DB 1; Length 742;  
 Best Local Similarity 45.3%; Pred. No. 1.1e-45;  
 Matches 163; Conservative 56; Mismatches 101; Indels 40; Gaps 9;

QY 7 IKVVVRPPNAREIDGAKCIVRMENGTI--LTTPPGAEEKARKSGKTMGPKAF 64  
 Db 9 VVVVVRPPNAREIDGAKCIVRMENGTI--LTTPPGAEEKARKSGKTMGPKAF 64

QY 65 DRSY-WSFQKNAPYARQEDFDGLGVPLLDNAFGYNNCFAYGOTSGSKYSNMGYCK 123  
 Db 59 DTVYDWN-----SKQILDVDETFRSLVESVLOGFNGTIFAYGOTGKFTFMESVRS 110  
 QY 124 E---HGVIPRICQDMFRINELQKDKMLCTVEYSLEYIYNERVRLNFTSKGNLAK 180  
 Db 111 NPELRGVIPNSFEHIFTIARTQNOQL---VRASYLEYIYOEETROLLAKDKDLQKLE 167  
 QY 181 HPSTGPVVEDLAKLVRSFOETENLMDGNKARTVAATNNNNTSSSRSHAVTLTITQKH 240  
 Db 168 PRTGVYVVDKSSPVTKSVLEHRTVMYVGNNSVSGSTNNNSRSHAFITITL-----E 223  
 QY 241 DEETKQTE---KVKISIVDLASGRATSGTGAIRKEGAENISLSTGLGWTAALAD 297  
 Db 224 CSELGVGENHIRGVKLNLDVLASGERQTKGTGAORLKEATKINLSLSTGLNVISSALVD 283  
 QY 298 MSSGQKKQLPVPRDYSVLTMLKDSLGNSMTAMTAISPADINFEETLSTLYRADSAK 357  
 Db 284 -----GKSSHPIYRNSKTLRLQDSLGNAKTVMWANMGPASYNDEETITLYRANRAK 337

## RESULT 12

A47334

LeKin kinesin-related antigen - Leishmania chagasi (fragment)

C: Species: Leishmania chagasi

C: Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 24-Sep-1999

C: Accession: A47334

R: Burns Jr., J.M.; Shreffler, W.G.; Benson, D.R.; Ghallib, H.W.; Badaro, R.; Reed, S.G

Proc. Natl. Acad. Sci. U.S.A. 90, 775-779, 1993

A: Title: Molecular characterization of a kinesin-related antigen of Leishmania chagasi

A: Reference number: A47334; MUID: 95133867

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-955 &lt;BUR&gt;

A: Cross-references: GB: I07879; NID: g308884; PIDN: AAA29254.1; PID: g308885

A: Experimental source: MHOM/BR/82/BA-2.C1

A: Note: sequence extracted from NCBI backbone (NCBI:122864, NCBI:122865)

C: Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology

C: Keywords: ATP; P-loop

F:132-398/Domain: kinesin motor domain homology &lt;KMT&gt;

F:122-129/Region: nucleotide-binding motif A (P-loop)

Query Match 38.5%; Score 705.5; DB 2; Length 955;  
 Best Local Similarity 42.7%; Pred. No. 5e-45;  
 Matches 163; Conservative 56; Mismatches 126; Indels 35; Gaps 8;

QY 7 IKVVVRPPNAREIDGAKCIVRMENGTI--LTTPPGAEEKARKSGKTMGPK 57

Db 13 VVVVVRPPNAREIDGAKCIVRMENGTI--LTTPPGAEEKARKSGKTMGPK 57

QY 58 GKPAFAEDTSYSGFNAPYARQEDFDGLGVPLLDNAFGYNNCFAYGOTSGSK 114

Db 73 ----FOEDHVTFVSTPDCAGATATQADVFRTGCVPLVQHAEDGFNSCLFAYGOTSGSK 128

QY 115 SYSNMG-----YGKEHGVIPRICQDMFRINELQKDKMLCTVEYSLEYIYNERVRL 168

Db 129 TYTMGADVSAISGEGNVTPRICLEIFETARKASVPAQGSRWIVELGVYEVYNERVRL 188

QY 169 NPSTGK-----NKVREHSTGTPVEDLAKLVRSFOETENLMDGNKARTVAATNM 220

Db 189 GKRRKGVKGGEVYVVDREHPSRGSFVGVGQRLVEVGSDDVRLIEIGNGVRHTASTKM 248

QY 221 NETSSSRSHAVTLTITQKH-----TQKHDEETKQTEKAKIISLVDLASERATSGTGA 276

Db 249 NDNRSRSHAVTLTITQKH-----TQKHDEETKQTEKAKIISLVDLASERATSGTGA 308

QY 277 EGAEENISLSTGLGWTAALADSS--GKQKKQNLVYPRDYSVLTMLKDSLGNSMTAMTA 335

Db 309 EATHINISLSTGLGWTAALADSS--GKQKKQNLVYPRDYSVLTMLKDSLGNSMTAMTA 368

QY 336 ISPADINFEETLSTLYRADSAK 357





F:100-107/Region: nucleotide-binding motif A (P-loop)  
F:369-599/Domain: helical rod #status predicted <ROD>  
F:600-701/Domain: tail globular #status predicted <TGL>  
F:106/binding site: ATP (Lys) #status predicted

```
Query Match          37.6%; Score 689; DB 1; Length 701;
Best Local Similarity 45.6%; Pred. No. 5.6e-44;
Matches 164; Conservative 50; Mismatches 106; Indels 38; Gaps 10;

Oy 6 NIKVVVVRFPNAREID--RGAKCTVVRGSGNOTILTPPGAEEKARKSGKTIWGPKAF 62
Db 14 NIKVVVVRCPLENEKSKMCPQAVSVDEMRCVITV-----HKTDSNN---EPPKTF 61

Oy 63 AFDRSYMSFDKNAPNVARQEDLPDGLGVPLLDNAFGYNNCIFAYGQTGSKSKSYNMGYG 122
Db 62 TFDVTF-----GPE-SKQLDVNLTARPIDSVLEGYNGTIFAYGQTGSKTFMESVR 114

Oy 123 KE---HGVIPRICDMFRINELQKDKLTCTVEVSYLEYNERNVRLDLPSTKGNLKYR 179
Db 115 AVPGLAGVTPNSFAHFPGHIAKAEQDTRP--LVRVSYLEYNEEVRDLGDKDQTORLEVY 172

Oy 180 EHPSTGPTVVEDLAKLVVRSFOETENLMDGCKNKARTVATRNMTSSSHAVFTLL--TQ 237
Db 173 EKPVGVTIKDLGAYVYNADMDRIHLGHKNSVGAHNHSSHRAIFITIECS 232

Oy 238 KWHDEETWDTQEKAKISLDLAGSFRATGATGABLKSGAEINSLSTLGRVIALAD 297
Db 233 KGVDKGNHV---RMGKLHLVDLAGSERQAKTGATGQRLKEATKINLSLSTLGNVISALVD 289

Oy 298 MSSGKKKKNOLVPYRDSVLTWLLKDSLGNSMTAMTAAISPADINPEETLSLTRYADSAK 357
Db 290 -----GKSTHVPYRNSKLTRELQDSLGNSKTMKCANIGPADYNYDETISTLRYANRAK 343
```

Search completed: April 25, 2001, 10:14:12  
Job time: 288 sec

**THIS PAGE BLANK (USPTO)**



```

FT DOMAIN 622 681 COILED COIL (POTENTIAL).
FT DOMAIN 801 823 COILED COIL (POTENTIAL).
FT DOMAIN 1575 1673 PH.
FT NP_BIND 97 104 ATP (POTENTIAL).
SQ SEQUENCE 1690 AA: 191083 MW: D8DDEC784624FB4D CRC64;

Query Match 62.8%; Score 1152; DB 1; Length 1690;
Best Local Similarity 62.1%; Pred. No. 1.3e-75;
Matches 226; Conservative 58; Mismatches 56; Indels 24; Gaps 6;

QY 4 GNIKVVVRPNAREIDRACIVRMENQITLTPPGABEKARKSGKTTMDGPKAPA 63
DB 3 GASYKVVAVRPNRSREMSRDSKCIQSGSTTTVNPKPQKET-----PKSPS 51
QY 64 FDRSYWSPDKNAP---NTARQEDLPDLGVPLLDNAKFGVNCIFAYGTGSGKSYMMG 120
DB 52 FDSYWS---HTSPEDINTASOKVVRDYGEMIQHAFEGVNCIFAYGTGAGKSYMMG 109
QY 121 YGK---EHGVIPRCDQMFRRINELQKDKTLCTVSVYLEITNERNVRLDNLNPKGLAV 178
DB 110 KQEKDOQGIIPQLCELESRLDNTND--NMSYSVEVSYMBIYCERVRDNLNPKKGLAV 168
QY 179 REHPSGPGYVEDLAKLVRSFOEINLMDGKNKARTVAATNMNETSSRSHAVFTLTOK 238
DB 169 REHPLGPGYVEDLSKLVTSYNDIQLMDSGNKARTVAATNMNETSSRSHAVFTLTOK 228
QY 239 WHDEETKMDTKVAKLSVLVDLAGSERATSTGATGARLKEGAEINRSLTIGVIAALADM 298
DB 229 RHDAETNITTEKYSKLSVLVDLAGSERADSTGANGTRLKEGANINKSLTIGKVISALAE 288
QY 299 SSG-----KQKKNLQVPRDSVLTMLKDSGNSMTAMTAISPADINFEETLSTLYA 353
DB 289 DSGPNKNNKKKTKDFIPYRDSVLTMLLRNLSGNSRTAMVAALSPADINDETSLTRYA 348
QY 354 DSAA 357
DB 349 DRAK 352

RESULT 2
ID KPIA_MOUSE STANDARD: PRT: 1695 AA.
AC 01-OCT-1993 (rel. 27, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DE 15-JUL-1998 (rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KPIA.
GN KPIA OR KIF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NC KPIA_TextID:10090;
RA KPIA_MOUSE
RC SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA MEDLINE=55292344; Pubmed=7539720;
RA Okada Y, Yanazaki H, Sekine-Aizawa Y, Hirokawa N;
RA "the neuron specific kinesin superfamily protein KPIA is a unique
RA monomeric motor for anterograde axonal transport of synaptic vesicle
RA precursors";
RA Cell 81:769-780(1995).
RA [2]
RP PRELIMINARY SEQUENCE OF 100-247 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=93077686; Pubmed=1447303;
RX Aizawa N., Sekine Y., Takemura R., Zhang Z., Nangaku M.,
RX Hirokawa N.;
RX "Kinesin family in murine central nervous system.";
RT J. Cell Biol. 119:1287-1296(1992).
CC -1- FUNCTION: MOTOR FOR ANTEROGRADE AXONAL TRANSPORT OF SYNAPTIC
CC VESICLE PRECURSORS.
CC -1- SUBUNIT: MONOMER.

```

```

-1- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN
TISSUE (MAINLY IN THE CEREBELLUM AND CEREBRUM) WITHIN A SINGLE
TYPE OF NEURONAL CELL. WITHIN THE NEURONAL CELL LEVELS ARE
CONCENTRATED AROUND THE AXON, WITH SMALLER AMOUNTS IN THE
PERINUCLEAR AND SYNAPTIC REGIONS.
-1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
SUBFAMILY.
-1- SIMILARITY: CONTAINS 1 PH DOMAIN.
-1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; D29951; BAA06221.1; -
DR PIR; E44259; E44259.
DR HSSP; P17119; 3KAR.
DR MGD; MGI:108391; Kifla.
DR InterPro; IPR000253; -.
DR InterPro; IPR001752; -.
DR InterPro; IPR001849; -.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR Motor protein; Microtubules; ATP-binding; Coiled coil.
FW DOMAIN 1 361
FT DOMAIN 366 383 MECHANOCHEMICAL (MOTOR).
FT DOMAIN 429 462 COILED COIL (POTENTIAL).
FT DOMAIN 516 572 COILED COIL (POTENTIAL).
FT DOMAIN 622 691 FHA.
FT DOMAIN 801 822 COILED COIL (POTENTIAL).
FT DOMAIN 1580 1678 COILED COIL (POTENTIAL).
FT NP_BIND 97 104 ATP (BY SIMILARITY).
SQ SEQUENCE 1695 AA: 191724 MW: D6EC3B88C9CC66 CRC64;

Query Match 62.5%; Score 1147; DB 1; Length 1695;
Best Local Similarity 61.8%; Pred. No. 3.1e-75;
Matches 225; Conservative 58; Mismatches 57; Indels 24; Gaps 6;

QY 4 GNIKVVVRPNAREIDRACIVRMENQITLTPPGABEKARKSGKTTMDGPKAPA 63
DB 3 GASYKVVAVRPNRSREMSRDSKCIQSGSTTTVNPKPQKET-----PKSPS 51
QY 64 FDRSYWSPDKNAP---NTARQEDLPDLGVPLLDNAKFGVNCIFAYGTGSGKSYMMG 120
DB 52 FDSYWS---HTSPEDINTASOKVVRDYGEMIQHAFEGVNCIFAYGTGAGKSYMMG 109
QY 121 YGK---EHGVIPRCDQMFRRINELQKDKTLCTVSVYLEITNERNVRLDNLNPKGLAV 178
DB 110 KQEKDOQGIIPQLCELESRLDNTND--NMSYSVEVSYMBIYCERVRDNLNPKKGLAV 168
QY 179 REHPSGPGYVEDLAKLVRSFOEINLMDGKNKARTVAATNMNETSSRSHAVFTLTOK 238
DB 169 REHPLGPGYVEDLSKLVTSYNDIQLMDSGNKARTVAATNMNETSSRSHAVFTLTOK 228
QY 239 WHDEETKMDTKVAKLSVLVDLAGSERATSTGATGARLKEGAEINRSLTIGVIAALADM 298
DB 229 RHDAETNITTEKYSKLSVLVDLAGSERADSTGANGTRLKEGANINKSLTIGKVISALAE 288
QY 299 SSG-----KQKKNLQVPRDSVLTMLKDSGNSMTAMTAISPADINFEETLSTLYA 353
DB 289 DSGPNKNNKKKTKDFIPYRDSVLTMLLRNLSGNSRTAMVAALSPADINDETSLTRYA 348

```

```

QY 354 DSAK 357
DB 349 DRAX 352

RESULT 3
KFIB_HUMAN STANDARD; PRT; 1103 AA.
AC Q43896; 075186; 13-DEC-1998 (rel. 37, Created)
DT 13-DEC-1998 (rel. 37, Last sequence update)
DE KINESIN-LIKE PROTEIN KIF1B.
GN KIF1B OR KIAA0706.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=98352063; PubMed=9685376;
RA Dornier C., Closek T., Mueller S., Moeller N.P.H., Ullrich A.,
RA Lamers R.;
RA Characterization of KIF1B, a new kinesin-like protein involved in
RT vesicle transport from the Golgi apparatus to the endoplasmic
RL reticulum;
RL J Biol. Chem. 273:20267-20275(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RA *Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.*;
RL DNA Res. 5:169-176(1998).
RN [3]
RP FUNCTION: MOTOR REQUIRED FOR THE RETROGRADE TRANSPORT OF GOLGI
CC VESICLES TO THE ENDOPLASMIC RETICULUM. HAS A MICROTUBULE PLUS END-
CC DIRECTED MOTILITY.
CC [1]
CC TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED, WITH MOST
CC ABUNDANT EXPRESSION IN HEART AND SKELETAL MUSCLE.
CC [2]
CC PTM: PHOSPHORYLATED ON TYROSINE.
CC [3]
CC SIMILARITY: CONTAINS 1 FHA DOMAIN.
CC [4]
CC SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC [5]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB014606; AK52117.1; -.
CC FTH: 603060; F0000253; -.
CC Incepero; F00001752; -.
CC Pfam: PF00498; FHA: 1.
CC Pfam: PF00225; Kinesin; 1.
CC PRINTS: PR00380; KINESINHA.
CC PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
CC PROSITE: PS50067; FHA_DOMAIN; FALSE NEG.
CC Motor protein. Microtubules; ATP-binding; Coiled coil;
CC phosphorylation.
CC [6]
CC MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
CC DOMAIN 1 355
CC COILED COIL (POTENTIAL).
CC DOMAIN 359 388
CC COILED COIL (POTENTIAL).
CC DOMAIN 438 479

QY 523 590
DB 633 674
DB 828 872
DB 97 104
DB 669 669
DB 955 962
DB 976 977
DB 1103 AA: 123071 MM; F14BC39BD4B2EC10 CRC64;

Query Match 61.3%; Score 1124; DB 1; Length 1103;
Best Local Similarity 61.3%; No. 8,4e-7;
Matches 221; Conservative 59; Mismatches 61; Indels 16; Gaps 5;

QY 4 GSNIVVVVVPFNARETDGAKCIYVMEGNOTILTPPGAEEKARKSKGTIMQPKAFA 63
DB 3 GASVAVVAVRFPNARETSQAKCVSMOGNTTSINP-----KOSDKAPKSPT 51
QY 64 FDRSYVDFKNA-PNTARQEDLFQDLGVPLLDNAFKYNNCTFAYQGTGCKSKYSMMGYG 122
DB 52 FDSYVSWSTSTEDPQASQQQYVYDICEEMLLHAFEGYVNCVFAYQGTGAKSKYTMHGQ 111
QY 123 K-EHGVIPRICQDMFRINELQKNNKTCFVSVSYLEIYNERVROLLNFSTGNKLVRE 180
DB 112 EPQCGIVPOLCEDLFSVSENG-SAQLSYSVSVSYMEIYCHRVROLLNPKRSGSLKVR 170
QY 181 HPTGPTVYDLAKIVRSFOETENLADGNKARTVAATNNNETSSSHAVFTLITOKWH 240
DB 171 HPLGPTVQVLSKLVSTYADLADLADGKAKVTAATNNNETSSSHAVFTIVTQCH 230
QY 241 DEETKMTQTEKVAKTSIYDLAGSERATSGATGARKLKGCAETNRSLSLTGLRVTAALADSS 300
DB 231 DQLTGLDSEKVSLSIYDLAGSERADSSGARGHRLKSGANINKSLTTLGKVISALADMS 290
QY 301 GKOKNOLVPRDVSMTWLLKLDLGGNSMTAAISPADINFEETLSTLRVADSAC 357
DB 291 -KKRKSDFIPYRDSVTWLLKENLGGNSMTAAISPADINVEETLSTLRVADRTK 346

RESULT 4
KFIB_HUMAN STANDARD; PRT; 1016 AA.
AC Q60333; 2000 (rel. 40, Created)
DT 01-OCT-2000 (rel. 40, Last sequence update)
DE KINESIN-LIKE PROTEIN KIF1B (KIFB).
DE KIF1B OR KIAA0591.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Bougueleret L., Dufaure-Gare I., Grel P.;
RT *DNA encoding a kinesin-like protein (hk1p) comprising biallelic
RT markers.*;
RL Patent number WO0063375, 26-OCT-2000.
RN [2]
RP SEQUENCE OF 479-1816 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RT *Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.*;
RL DNA Res. 5:31-39(1998).
RN [3]
RP SEQUENCE OF 1449-1816 FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Pujil A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,

```

RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;  
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RL -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A  
 CC MICROTUBULE PLUS END-DIRECTED MOTILITY (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND  
 CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AX039604; CAC16629.1; -;  
 CC EMBL; AB011163; BAA25517.1; -;  
 CC EMBL; AK022977; BAB14341.1; -;  
 CC InterPro: IPR000253; -;  
 CC InterPro: IPR001849; -;  
 CC Pfam: PF00169; PH; 1;  
 CC Pfam: PF00498; FHA; 1;  
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1;  
 CC PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1;  
 CC PROSITE; PS00606; FHA DOMAIN; 1;  
 CC PROSITE; PS00003; PH DOMAIN; 1;  
 CC Motor protein; Microtubules; ATP-binding; Coiled coil.  
 KW DOMAIN 1 361 MECHANOCHEMICAL (MOTOR).  
 FT DOMAIN 365 486 COILED COIL (POTENTIAL).  
 FT DOMAIN 470 502 COILED COIL (POTENTIAL).  
 FT DOMAIN 556 612 FHA.  
 FT DOMAIN 668 737 COILED COIL (POTENTIAL).  
 FT DOMAIN 841 869 COILED COIL (POTENTIAL).  
 FT DOMAIN 1702 1799 PH.  
 FT NP\_BIND 97 104 ATP (POTENTIAL).  
 SQ SEQUENCE 1816 AA; 204474 MW; AD62F0515978C783 CRC64;

Query Match 61.08; Score 1119; DB 1; Length 1816;  
 Best Local Similarity 60.88; Pred. No. 3.7e-73;  
 Matches 220; Conservative 59; Mismatches 63; Indels 20; Gaps 5;  
 QY 4 GNIKVVYVPEPAREIDRAGCAKIVYMGNOTILTPPGAEKARKSKITMGDKPAPA 63  
 DB 3 GASKVAVVAPRPSRETSKSCILIQNGSTSTINKPKE-----AKPSFS 51  
 QY 64 FDRSYNSP-DKNAPYARQEDLPDLGVFLIDNKAQYKYNICFAYGTGSKSYSMGVG 122  
 DB 52 FDSYNSHTSPEDCFASQNRVYNDIGKMLHAEQYVNCIFAYGTGAGKSTYMGKQ 111  
 QY 123 KEH--GVIPTRICODMPRELQKDKNLCTGVESYLEIYNERVRLDLPSTKGNLKYVE 180  
 DB 112 EESQAGIIPOLCEELFERIND-NCNEEMSYSEVSYMEIYCVSRVOLLNPKNKNLRVE 170  
 QY 181 HPSTGPGYVEDLAKLVVRSFQRIENLMDGKNKARTVAATNNETSSRSRSHAVFTLTQKH 240  
 DB 171 HPLGLPGYVEDLSKLVATSYTDIADLMDAGNKARTVAATNNETSSRSRSHAVFTLTQKH 230  
 QY 241 DEETKMDYTKVAKISLVDLAGSERATGTGATGARLKEGAFENRSLSTLGRVTAALADM-- 298  
 DB 231 DNETNLSTERYKYSILVDLAGSERADSTGAKTRLKEGANINKSLTLGKVISALAEVDN 290  
 QY 299 ---SSGKKKQNLQVPRDSVLTWLLKSLGNGSNMTAMIAAISPADINFETILSTLAYADS 355  
 DB 291 CTEKSKKKKKTKDFIPYRDSVLTWLLRNLGNSRTAMVAALSPADINDETILSTLAYADR 350  
 QY 356 AK 357

Db 351 AK 352  
 II  
 RESULT 5  
 KF1B\_MOUSE STANDARD; PRT; 1816 AA.  
 AC Q60575; Q9WVE5; Q9R0B4; Q9Z119;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE KINESIN-LIKE PROTEIN KIF1B.  
 GN KIF1B.  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Brain;  
 RX MEDLINE=95094296; PubMed=7528108;  
 RA Yamagaki M., Sato-Yoshitake R., Okada Y., Noda Y., Takemura R.,  
 RA "KIF1B, a novel microtubule plus end-directed monomeric motor protein  
 RT for transport of mitochondria.";  
 RL Cell 79:1209-1220(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=ICR; TISSUE=Brain;  
 RX PubMed=10571041;  
 RA Gong T.L., Winnicki R.S., Kohman D.C., Lomax M.I.;  
 RT "A novel kinesin of the UNC-104/KIF1 subfamily encoded by the Kif1b  
 RT gene.";  
 RL Mamm. Genome 10:0-0(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RA Conforti L., Buckmaster A., Tarlton A., Brown M.C., Lyon M.F.,  
 RA Perry V.H., Coleman M.P.;  
 RT "The major brain isoform of Kif1b lacks the putative mitochondria-  
 RT binding domain.";  
 RL Mamm. Genome 10:0-0(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=ICR;  
 RA Nakagawa T., Hirokawa N.;  
 RT "Identification and characterization of a new kinesin superfamily  
 RT KIF1B-beta.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: MOTOR FOR ANTEROGRADE TRANSPORT OF MITOCHONDRIA. HAS A  
 CC MICROTUBULE PLUS END-DIRECTED MOTILITY.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: PREDOMINANTLY FOUND ON THE VESICLE- AND  
 CC TUBULE-LIKE STRUCTURES WITHIN THE CYTOPLASM.  
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2/BETA AND 3;  
 CC ADDITIONAL ALTERNATIVE PRODUCTS: 122 SPLICING.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED ALMOST EXCLUSIVELY IN ADULT BRAIN  
 CC TISSUE MATURING IN THE CEREBELLUM AND CEREBRUM WITHIN A SINGLE  
 CC TYPE OF NEURONAL CELL.  
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 FHA DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D17577; BAA04503.1; -;



```

FT DOMAIN 777 797 COILED COIL (POTENTIAL).
FT DOMAIN 183 335 MICROTUBULE-BINDING.
FT DOMAIN 1460 1558 PH.
FT DOMAIN 957 1052 ARG/LYS-RICH (BASIC).
FT DOMAIN 1203 1584 ARG/LYS-RICH (BASIC).
FT NP_BIND 93 100 ATP (BY SIMILARITY).
FT VARIANT 598 598 I -> T.
FT VARIANT 930 930 V -> M.
SQ SEQUENCE 1584 AA; 179736 MW; 559462B3FD029B43 CRC64;

Query Match
Best Local Similarity 61.2%; Score 1106.5; DB 1; Length 1584;
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

QY 6 NIKVYVRPNAREIDRGAKCIVRMENQNTILTPPGAEKARKSGKTIIDGPKAFAD 65
DB 3 SVKAVYVRPNAREISNKTGCVLQVNGNTT-----NGHSINKENKSFNFD 50

QY 66 RSYWSFDKNAFYARQEDLDQGLVPLDPAFGKYNVNCIFAYGQSGSKSYMGGYK 123
DB 51 HSYWSFARNDFHTTKQVVEELGVEMLEHAFEGYNVCIFAYGQSGSKSYMGGKANDP 110

QY 124 -EHSVPRICQDMFRINELQDKNLCTVEVSYLEIYNERVRLDLPSTKGNLKVREHP 182
DB 111 DEMGIIPRLCNDFARIDN-NNDKQVQSVSEYMEIYCYERVKDLLNPSGNGNLVREHP 169

QY 183 STGPTVEDLAKLVRSFOEINLMDGKARTVAATNMNNTSSRSHAVFTLTQKHDE 242
DB 170 LLGYPVDLTKMAVCSDHCNLMDEGKARTVAATNMNNTSSRSHAVFTLTQKHCA 229

QY 243 ETRMDTEKVKISLVDLAGSERATSGATGARLKEGAEINRSLTGLRVIAALADMSCK 302
DB 230 DSNLDTEKHSKISLVDLAGSERANTSGEQRKEGANKSLTGLVLSKLAESTKK 289

QY 303 OKNQ-LVPPYSDSVLWLLKDSLGNSMTAMTAASPADINFEETLTSLRYADSAX 357
DB 290 KSKNKGVPYSDSVLWLLHNGJGNSKTMALAAASPADINFEETLTSLRYADRAK 345

RESULT 7
ID KPID_RAT STANDARD; PRT; 1097 AA.
AC Q3578; 2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DR KINESIN-LIKE PROTEIN KIF16.
GN KIF16.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RP STRALIN-HISTAR.
RA Rogers K.B., Griffin M., Brophy P.J.;
RT "The secretory epithelial cells of the choroid plexus employ a novel
RT kinesin-related protein".
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTRAINS 1 FHA DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use the European Bioinformatics Institute. There are no restrictions on its
CC use and profit institutions as long as its content is in no way
CC modified and this statement is prominently displayed, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; AJ000696; CAA04248.1;

```

```

DR HSP: P17119; 3KAR.
DR InterPro: IPR000253; -.
DR InterPro: IPR001752; -.
DR Pfam: PF00225; Kinesin; 1.
DR Pfam: PF00498; FHA; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR PROSITE: PS00411; KINESIN MOTOR DOMAIN1; FALSE_NEG.
DR PROSITE: PS50067; KINESIN MOTOR DOMAIN2; 1.
DR PROSITE: PS50068; FHA DOMAIN; FALSE_NEG.
DR Motif protein; Microtubules; ATP-binding; Coiled coil.
KW DOMAIN 1 357 MECHANOCHEMICAL (MOTOR).
FT DOMAIN 358 380 COILED COIL (POTENTIAL).
FT DOMAIN 437 478 COILED COIL (POTENTIAL).
FT DOMAIN 520 587 FHA.
FT DOMAIN 630 671 COILED COIL (POTENTIAL).
FT DOMAIN 824 868 COILED COIL (POTENTIAL).
FT NP_BIND 96 103 ATP (POTENTIAL).
SQ SEQUENCE 1097 AA; 122333 MW; 8EF40B1C7579BA5B CRC64;

Query Match
Best Local Similarity 60.5%; Score 1085.5; DB 1; Length 1097;
Matches 216; Conservative 64; Mismatches 60; Indels 17; Gaps 7;

QY 4 GGNIKVYVRPNAREIDRGAKCIVRMENQNTILTPPGAEKARKSGKTIIDGPKAF 63
DB 3 GASVYVVRPNARETSQDAKCVYSMQGNTTINP-----KQSRMFL---KA-S 50

QY 64 FDSYSWFDK-NAPNTARQEDLDQGLVPLDPAFGKYNVNCIFAYGQSGSKSYMGGYK 122
DB 51 FDSYSWHSYTSDEPQFASQQVYRDIQGEMLLHAFEGYNVCIFAYGQSGSKSYMGGQ 110

QY 123 K-EHGVPRICQDMFRINELQDKNLCTVEVSYLEIYNERVRLDLPSTKGNLKVRE 180
DB 111 PQQQIYVQQLCEDLFSRVN-VNQASLYSVSEYMEIYCYERVKDLLNPKSRGSLVRE 169

QY 181 HPTSGPYVEDLAKLVRSFOEINLMDGKARTVAATNMNNTSSRSHAVFTLTQKHV 240
DB 170 HPILGYPVDLTKMAVTSYADIADMGKARTVAATNMNNTSSRSHAVFTVTFQSH 229

QY 241 DEETKMDTEKVKISLVDLAGSERATSGATGARLKEGAEINRSLTGLRVIAALADMS 300
DB 230 DLTGLDSEKVKISLVLNLAGSERADSSGARGRLKKEGANKINSLTGLKVISALADLS 289

QY 301 GKQKQLVYPYSDSVLWLLKDSLGNSMTAMTAASPADINFEETLTSLRYADSAX 357
DB 290 -KRRKSDFTIPROSVLWLLKNGLSRNTAMTAAASPADINFEETLTSLRYADTKR 345

RESULT 8
ID KPID_RAT STANDARD; PRT; 689 AA.
AC Q8656; 2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DR KINESIN-LIKE PROTEIN KIF18 (FRAGMENT).
GN KIF18.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RP Faure K., Gruber D., Bulinski J.C.;
RT Identification of kinesin-like molecules in myogenic cells.*;
RL Eur. J. Cell Biol. 0:0-0(1998).
CC -1- FUNCTION: MOTOR FOR ANTIGRADE TRANSPORT OF MITOCHONDRIA. HAS A
CC -1- CYTOSOLIC PLUS END-DIRECTED MOTILITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PRESONARIA CYTOSOL; ON THE VESICLE- AND
CC -1- SUBCELLULAR STRUCTURES WITHIN THE CYTOSOL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. UNC-104
CC SUBFAMILY.

```



OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinolidea; Euechinolidea; Echinacea; Echinoida; Strongylocentrotidae;

[illegible]

Best Local Similarity	46.0%	Pred. No. 1.7e-46;	
Matches 167:	Conservative	62;	Mismatches 99;
			Indels 35;
			Gaps 10;

[illegible]

```

RESULT 10
ID K121_STRPU STANDARD; PRT; 742 AA.
AC P46874;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE KINESIN-II 95 KDA SUBUNIT (KRP-85/95 95 KDA SUBUNIT).
GN KRP95.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
NCBI_Taxid=7668;
RN [1]
RX "Novel heterotrimeric kinesin-related protein purified from sea
RT urchin eggs."
RL Nature 366:268-270(1993).
CC -1- SUBUNIT: TRIMER OF A 115 KDA SUBUNIT AND TWO KINESIN-LIKE SUBUNITS
CC OF 95 AND 85 KDA.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00936; AAA87393.1; -.
DR HSSP; P56536; 2KIN.
DR InterPro; IPR001752; -.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00360; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR MOTOR protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 338 337 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 618 943 COILED COIL (BY SIMILARITY).
FT DOMAIN 95 102 GLOBULAR (BY SIMILARITY).
FT NP_BIND 95 102 ATP (POTENTIAL).
FT SEQUENCE 742 AA; 84202 MW; 47C40A367BA47B5 CRC64;
DR -----
Query Match 38.98; Score 712; DB 1; Length 742;
Best Local Similarity 45.38; Pred. No. 3,3c-44;
Matches 163; Conservative 56; Mismatches 101; Indels 40; Gaps 9;
QY 7 LKVVVRVPPNAREIDRCACIVRMGNOTI--LPNPPGAEEKARKSGKTIMDGKPAFAF 64
DB 7 LKVVVRVPPNAREIDRCACIVRMGNOTI--LPNPPGAEEKARKSGKTIMDGKPAFAF 64
FT REPEAT 618 943
FT REPEAT 899 939
FT REPEAT 938 955
FT NON_TER 955
QY 65 DRSY-HSFDKAPNVARQEDLPDGLGVLDNAPKYNNCIFAYGQTGSGKSYSMGCK 123
DB 65 DRSY-HSFDKAPNVARQEDLPDGLGVLDNAPKYNNCIFAYGQTGSGKSYSMGCK 123
QY 59 DTVTDNN-----SKQIDLYDFRSLVESVLOGENGTFIFAYGQTGSGKTIFMEGVRS 110
DB 59 DTVTDNN-----SKQIDLYDFRSLVESVLOGENGTFIFAYGQTGSGKTIFMEGVRS 110
QY 124 E---HGVIPRICDFRNELQKNTCTVEVSELYEINERNVRDLNPSTKGNLKVRE 180
DB 124 E---HGVIPRICDFRNELQKNTCTVEVSELYEINERNVRDLNPSTKGNLKVRE 180
DB 111 NPRLRCVIPSFEHFTHARTNQOQFI--VRASYLEIVQEEIRDLAKDQKRLDLAE 167
QY 181 HPSTGPIYEDLAKLVVRFQBIENLMDGKNKARTVAATNNNTSSRSRAVFTLTLTKWH 240
DB 181 HPSTGPIYEDLAKLVVRFQBIENLMDGKNKARTVAATNNNTSSRSRAVFTLTLTKWH 240
DB 168 RPDFTGVYKDLSSFVTSKVEIKHVTYVGNRRSVGVTNNNCHSRSHAFITFI-----E 223
QY 241 DEETKMDTE--KVAIKSLVDLAGSRATSTGATGARLKEAGIENRSLSTLGRVIAALAD 297

```

```

DB 224 CSELGVDGENHVRGKLNLDVLAGSERQAKTGATGDRLKATKLNLSLGAUNVISALVD 283
QY 298 MSSCKOKKNOIYPRDSVPTWLTKDSLGSGNSMTAMIAISPADINFEETLSLRYADSAK 357
DB 284 -----GRSSHIPRDSKLPRLQLQDSIGGNAKTVVWVMGSPASYNFDETTITLRYANRAK 337
RESULT 11
KINL_LEICH
ID KINL_LEICH STANDARD; PRT; 955 AA.
AC P46865;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE KINESIN-LIKE PROTEIN K39 (FRAGMENT).
GN KIN.
OS Leishmania chagasi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
NCBI_Taxid=5669;
RN [1]
RX SEQUENCE FROM N.A.
RC STRAIN=MHON/BR/82 / ISOLATE BA-2;
RX MEDLINE=93133867; PubMed=8421715;
RX Burns J.M. Jr., Shreffler W.G., Benson D.R., Ghallib H.W., Badaro R.,
RA Reed S.G.;
RT "Molecular characterization of a kinesin-related antigen of
RT Leishmania chagasi that detects specific antibody in African and
RT American visceral leishmaniasis."
RL Proc. Natl. Acad. Sci. U.S.A. 90:775-779(1993).
CC -1- DEVELOPMENTAL STAGE: PREDOMINANT IN AMASTIGOTES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07879; AAA29254.1; -.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; -.
DR Pfam; PF00225; Kinesin; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
DR MOTOR protein; Microtubules; ATP-binding; Coiled coil; Repeat.
FT DOMAIN 1 399 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT DOMAIN 426 >955 COILED COIL (POTENTIAL).
FT NP_BIND 122 129 ATP (POTENTIAL).
FT DOMAIN 704 >955
FT REPEAT 704 742
FT REPEAT 743 781
FT REPEAT 782 820
FT REPEAT 821 859
FT REPEAT 860 899
FT REPEAT 898 937
FT REPEAT 938 >955
FT NON_TER 955
FT SEQUENCE 955 AA; 106168 MW; 8CA76815B884C6B9 CRC64;
DR -----
Query Match 38.58; Score 705.5; DB 1; Length 955;
Best Local Similarity 42.78; Pred. No. 1,4e-43;
Matches 163; Conservative 58; Mismatches 126; Indels 35; Gaps 8;
QY 7 LKVVVRVPPNARE--IDRCACIVRMGNOTILTPP-----PGAEKARKSGKTIMD 57
DB 7 LKVVVRVPPNARE--IDRCACIVRMGNOTILTPP-----PGAEKARKSGKTIMD 57
QY 13 VKVSVRVRPLNRENNAPEGTKVTVAQAQAAVTVKVLGSSNNGSAAESMGSTARRVAD 72
DB 13 VKVSVRVRPLNRENNAPEGTKVTVAQAQAAVTVKVLGSSNNGSAAESMGSTARRVAD 72
QY 58 GKPAFAFRDSRYWGF--DKNAPNYAQEDLPDGLGVLDNAPKYNNCIFAYGQTGSK 114
DB 58 GKPAFAFRDSRYWGF--DKNAPNYAQEDLPDGLGVLDNAPKYNNCIFAYGQTGSK 114

```

CC	EMBL; AB002357; BAA20815.1; -
DR	HSSP; P56336; 2KIN.
DR	MIM; 603754; -
DR	InterPro; IPR001752; -
DR	Pfam; PF00225; kinesin; 1.
DR	PRINTS; PR00380; KINESINHEAVY.
DR	PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR	Motor protein; Microtubules; ATP-binding; Coiled coil; Neurene.
KW	DOMAIN 1 345
FT	DOMAIN 346 579
FT	DOMAIN 580 747
FT	DOMAIN
FT	MECHANOCHEMICAL MOTOR (BY SIMILARITY).
FT	COILED COIL (BY SIMILARITY).
FT	GLOBULAR.

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: D26077; BAA05070.1; -.  
DR HSSP: P56536; 2KIN.  
DR MGD: MGI:107688; Kif3b.  
DR InterPro: IPR001752; -.  
DR Pfam: PF00225; Kinesin; 1.  
DR PRINTS: P00380; KINESINHEAVY.  
DR PROSITE: PS00411; KINESIN MOTOR\_DOMAIN1; 1.  
DR PROSITE: PS0067; KINESIN MOTOR\_DOMAIN2; 1.  
KW Motor protein; Microtubules; ATP-binding; Coiled coil; Neurone.  
FT DOMAIN 1 345 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
FT DOMAIN 346 579 COILED COIL (BY SIMILARITY).  
FT DOMAIN 580 747 GLOBULAR.  
FT NP\_BIND 96 103 ATP (POTENTIAL).  
FT DOMAIN 386 393 POLY-GLY.  
FT DOMAIN 394 405 POLY-GLU.  
FT DOMAIN 723 730 POLY-SER.  
SQ SEQUENCE 747 AA; 85288 MW; FA359AA190ECB47 CRC64;

Query Match 38.2%; Score 701; DB 1; Length 747;  
Best Local Similarity 44.6%; Pred. No. 2.2e-43;  
Matches 162; Conservative 57; Mismatches 100; Indels 44; Gaps 10;

QY 6 NIKVVVVRFPNARE----IDRGAKCIVRMGNQITLTTPPGAEKARKSGRTIMDGPKA 61

DB 9 SVRVVVRCPNMGKEKAASDKYVDVVKL--GVSVKPKG-----TSHMPKT 56

QY 62 FAFDRSY-WSPDKNAPFYAREDLFDGLGVPLDPAFKYNNCIFYAGQTSGSKSSMAG 120

DB 57 FTEDAYDWN-----AKQELDYETFPFLVDSVLQGFNGCTIFYAGQTGTGKTYTMEG 108

QY 121 Y--GKEGVIPRICQDMFRINELQKKNLCTVEVSYLEYINERVROLLNPSTGKLNK 177

DB 109 VRGDPKPKGVIPNDFSHFTHIRSQNOQYL---VRASYLEYIOEIRQLLSKDOTKRL 165

QY 178 VREHPSTGPVYEDLAKLVRSFOEINLMDGNKARTVAATNNNTSSSHAVFTLTQ 237

DB 166 LKRPDTGVYVKDLSFVTSKVEIEHVMMNQNRSGVATNNNEHSSHAIFVTI-- 223

QY 238 KWHDETKMDE---KVAKISLVDLAGSERATSTGATGARKKEAGINRSLTGRVIAA 294

DB 224 --ECSEVLGDEHNRVGLKLVLDLAGSERQAKTGAGGERLKATKINLSALGNVISA 281

QY 295 LADMSSGQKKNOILVPYDSDVLIWLLKDSLGNSMTAAIASPADINFEETLTIRYAD 354

DB 282 LVD-----GASTHIFPADSKLRLQLDSLGGNKNATVKNVANGPASTINVEETLTIRYAN 335

QY 355 SAK 357

DB 336 RAK 338

RESULT 14  
FL10\_CHLRE STANDARD; PRT; 786 AA.  
AC P46869;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

GN KINESIN-LIKE PROTEIN FLA10 (KHP1 PROTEIN).

CC Chlamydomonas reinhardtii.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

CC Chlamydomonadaceae; Chlamydomonas.

OK NCBI\_TaxID=3055;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=137;

RY MEDLINE:94299638; PubMed:8027176;  
RA Walther Z, Vashishtha M, Hall J.L.;  
RT "The Chlamydomonas FLA10 gene encodes a novel kinesin-homologous  
protein".  
RL J. Cell Biol. 126:175-188 (1994).  
CC 1- FUNCTION: PROBABLY INVOLVED IN FLAGELLAR ASSEMBLY AND MAINTENANCE.  
CC MAY PLAY A ROLE IN FLAGELLAR SYNTHESIS.  
CC 1- TISSUE SPECIFICITY: FLAGELLAR AXONEME.  
CC 1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN  
CC II SUBFAMILY  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L33697; AAA21738.1; -.  
DR HSSP: P17119; 3KAR.  
DR InterPro: IPR001752; -.  
DR Pfam: PF00225; Kinesin; 1.

DR PRINTS: P00380; KINESINHEAVY.

DR PROSITE: PS00411; KINESIN MOTOR\_DOMAIN1; 1.

DR PROSITE: PS0067; KINESIN MOTOR\_DOMAIN2; 1.

KW Motor protein; Microtubules; ATP-binding; Coiled coil.

FT DOMAIN 1 358 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).

FT DOMAIN 367 687 COILED COIL (POTENTIAL).

FT NP\_BIND 688 786 GLOBULAR (POTENTIAL).

FT DOMAIN 97 104 ATP (POTENTIAL).

FT DOMAIN 388 391 POLY-GLY.

FT DOMAIN 705 714 POLY-GLY.

FT DOMAIN 756 759 POLY-ASP.

SQ SEQUENCE 786 AA; 86671 MW; F90969203EB79F1B CRC64;

Query Match 37.8%; Score 694; DB 1; Length 786;

Best Local Similarity 43.5%; Pred. No. 7.5e-43;

Matches 163; Conservative 56; Mismatches 110; Indels 46; Gaps 9;

QY 2 SGGG--NIKVVVVRFPNAREIDRGAKCIVRME---GNQITLTTPPGAEKARKSGRTIM 56

DB 4 AGGSGSEKVVVVRCPNMGKEKADGRSIVDMVDAGOVKNRPNKADASEP----- 54

QY 57 DGKRAFADFDSY-WSPDKNAPFYAREDLFDGLGVPLDPAFKYNNCIFYAGQTSGSKS 115

DB 55 --FPAFTFDQVYDWN-----CQORDVFDITARPLDSCIESYNGTIFYAGQTSGKS 104

QY 116 YSMNGYKGE---HGVIPRICQDMFRINELQKKNLCTVEVSYLEYINERVROLLNPST 172

DB 105 HTMEGDKGPELPGLLIPNFRVYFELIARDSGTKEF--LVRSSYLEIYNEVRQLLGKH 162

QY 173 KGNLVREHPSTGPVYEDLAKLVRSFOEINLMDGNKARTVAATNNNTSSSHAVFT 232

DB 163 SKMKELKSPDRGVYVKDLSFVCKNVEEMKNVLLAGKNRQGVATLMQDSSKSHSIFT 222

QY 233 LTL-----TKWHDDETKMDETKVAKISLVDLAGSERATSTGATGARKKEAGIN 282

DB 223 ITTICEIEKLESAAQPKAKDDSNHVRVGLKLVLDLAGSERQAKTGATGDRLEKGIN 282

QY 283 RSLSTLGRVITAAADMSSGQKKNOILVPYDSDVLIWLLKDSLGNSMTAAIASPADIN 342

DB 283 LUSALGNVISALVQDGKSGH-----IPYDKSKLRLQLDSLGGNKTIVNANGPADWN 336

QY 343 FEETLSTIRYADSAS 357

DB 337 YDETSTLTIRYANRAK 351

RESULT 15

RF3A\_MOUSE

ID RF3A\_MOUSE

STANDARD; PRT; 701 AA.

173	ERPDYGVYIKLDSLVYNNADNDORIKTLGIHKVSEVGCATNKNWELSSRSHAFITTECS2	Db
238	KWHDSEETKMDTEKAKISLVLDLSPRASRATISGTATGARLKEGAEINRSLSLTGRLVIAALD	QY
238	KWHDSEETKMDTEKAKISLVLDLSPRASRATISGTATGARLKEGAEINRSLSLTGRLVIAALD	237
233	KGVDGNQK-----RAGLRLHSLVDLWDSRQKATGATGRLEKAKINLSLSTGRLVIAALD	QY
238	MSGSGGKHKVLLQVPEVYTWLLDSLGNSKATNATSPADINFEETLSLTGRLVIAALD	239
238	MSGSGGKHKVLLQVPEVYTWLLDSLGNSKATNATSPADINFEETLSLTGRLVIAALD	285
290	-----CKSHVYPPNSKIFRLIADSLGNSKTPMCANTGADYNYDETLSLTGRLVIAALD	343

	Query Match	37.6%	Score 689	DB 1	Length 701	
	Best Local Similarity	45.6%	Pred. No. 1.3e+42			
	Matches 164	Conservative 10	Mismatches 148	Indels 38	Gaps 10	
OY	6	NTKKVVRPPENAREID---EKAQCTVMEGSGNQTLLTPPGCAEKAKSGKVTMGCKAF	62			
DB	14	NKKVVRCPRLNEREKSKVCRQAVSDMEKTVV-----HKTDSN-----EPRFT	61			
OY	63	ADFSYKSDANAPYAREQDLFDGLVPLLDNFAFGVNCIPFAGCTGSGSKYSVMWG	122			
DB	62	TDFTVP-----GPE-SKQLDVNLATRLPIIDSVLESGYNGTIFAGGCTGCTKTFMECV	114			
OY	123	KE---HWGIVPICOMDFRI NELQDKKULCTVSEVSYLETIVNRRVDDLNSKGNLKV	179			
DB	115	AVPGLRGVTPSPFAHTGHIKAEGDTRP--LVRSYLETIVNEEVRDLGKQDQTOLEVK	172			
OY	180	EPHSTGPVYEDLAKLVRSFOETNIDEGCNKATYVAATNNKSSSRSHAVFLTL--TG	237			



Result	No.	Score	Query	Match	Length	DB	ID	Description
1	1124	61.3	1103	4	075186			O75186 homo sapien
2	1114	60.7	1770	11	O92119			O92119 mus musculus
3	1110	60.5	1816	11	O9R0B4			O9R0B4 mus musculus
4	1109	60.4	1521	5	O9R811			O9R811 drosophila
5	1105	60.4	1521	5	O9R811			O9R811 drosophila
6	1106.5	60.3	1773	5	O9V766			O9V766 caenorhabdi
7	1102	60.1	1816	11	O9W6E5			O9W6E5 mus musculus
8	1085.5	59.2	1097	11	O35787			O35787 rattus norv
9	1081	58.9	2205	5	O9MG02			O9MG02 dictyosteli
10	1073.5	58.5	689	11	O88658			O88658 rattus norv
11	1028	56.1	1921	5	O01349			O01349 drosophila
12	1028	56.1	1921	5	O01349			O01349 drosophila
13	1016.5	55.4	1826	4	O9N078			O9N078 homo sapien
14	1016.5	55.4	1826	4	O9N078			O9N078 caenorhabdi
15	907.5	42.5	1316	2	O61039			O61039 drosophila
16	878.5	47.9	928	5	O09597			O09597 caenorhabdi
17	873	47.6	1174	5	O9VB25			O9VB25 drosophila
18	834	45.5	1648	4	O15058			O15058 homo sapien
19	793.5	43.3	1121	5	O9V1P4			O9V1P4 drosophila

```

Query Watch      61.3%  Score 1124;  DB 4;  Length 1103;
Best Local Similarity 61.3%  Prod NO. 2 3e-74
Matches 221; Conservative 59; Mismatches 61; Indels 16; Gaps 5;

QY 4 GNTTGVVVVRPNAREIDRGAKTIVMEGNQITLTTPPGABEKARKSKCTIMDGKPAFA 63
      :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

```

Db 3 GASVKVAVRVPFNARETSQDAKVVVSQGNQNTTSIINP-----KQSKDAKPSFT 51
Qy 64 FDRSYNSFDKNA-PNYARQEDLPDGLVPLLDNAFKYNNCIFAYGOTSGKSYSMGVG 122
Db 52 FDSYNSHTSTEDPQAFSQQVYRDIQGEMLLHAFEGYNYVCIFAYGOTGAGKSYTMGRQ 111
Qy 123 K--EHWGIPRICQDMFRINELQKDNCTVTVSVYLEIYNERNVRLDLPNPKGNLKVRE 180
Db 112 EPGQOQIVPQCELDLFSVSENO-SAQLSYSVEVSYMEIYCERVDLLNPKSRGSLRVE 170
Qy 181 HPSTGTYVEDLAKLVVRSFOEINLMDGKNKARTVAATNNNETSSRSHAVFTLTQKWH 240
Db 171 HPILGTYVDQLSKLAVTSYADIADLMDGKNKARTVAATNNNETSSRSHAVFTVFTQK 230
Qy 241 DEETKMDTEKVAKISLDVLAGSERATSTGATGARLKEGAEINRSLSLGRVTAALADMS 300
Db 231 DQTLGLDSEKYSISLDVLAGSERADSSGARGNRLKGCANINKSLTLTKGVISALADMS 290
Qy 301 GKQKKNQLVPPYRVSVLTLWLLKDSIGNSMTAMTAISPADINFEETLSLTRYADSK 357
Db 291 -KKKSDFTPYRDSVLTLWLLKENLGNSRTAMTAALSPADINVEETLSLTRYADRTK 346
RESULT 2
Q92119 PRELIMINARY; PRT: 1770 AA.
AC Q92119;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE KIFIB-BETA.
DE KIFIB-BETA.
GN KIFIB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakagawa T., Hirokawa N.;
RT Identification and characterization of a new kinesin superfamily
RL KIFIB-beta. (Feb-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: F027636.1; BAA5243.1; -.
DR HSRP; P33176.1; B62.
DR INTERPRO: IPR001849; -.
DR INTERPRO: IPR001752; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; Kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor cytoskeleton; Microtubules; ATP-binding; Coiled coil.
SQ SEQUENCE 1770 AA; 138850 MW; E444228501672E19 CRC64;

```

Query Match 60.7%; Score 1114; DB 11; Length 1770;  
 Best Local Similarity 61.3%; Pred. No. 2.5e-73;  
 Matches 219; Conservative 59; Mismatches 63; Indels 16; Gaps 5;

```

Qy 4 GGNKVVVVRVPFNAREIDRGAKCIVRMGQNTILTPPGAEEKARKSKTMDGPKAFA 63
Db 3 GASVKVAVRVPFNARETSQDAKVVVSQGNQNTTSIINPKE-----AKPSFS 51
Qy 64 FDRSYNSFDKNA-PNYARQEDLPDGLVPLLDNAFKYNNCIFAYGOTSGKSYSMGVG 122
Db 52 FDSYNSHTSTEDPQAFSQQVYRDIQGEMLLHAFEGYNYVCIFAYGOTGAGKSYTMGRQ 111
Qy 123 K--EHWGIPRICQDMFRINELQKDNCTVTVSVYLEIYNERNVRLDLPNPKGNLKVRE 180
Db 112 EPGQOQIVPQCELDLFSVSENO-SAQLSYSVEVSYMEIYCERVDLLNPKSRGSLRVE 170
Qy 181 HPSTGTYVEDLAKLVVRSFOEINLMDGKNKARTVAATNNNETSSRSHAVFTLTQKWH 240
Db 171 HPILGTYVDQLSKLAVTSYADIADLMDGKNKARTVAATNNNETSSRSHAVFTVFTQK 230

```

```

Db 112 EESQAGIIPQCELFEXIND-NCNEEASYSVEVSYMEIYCERVDLLNPKGNLKVRE 170
Qy 181 HPSTGTYVEDLAKLVVRSFOEINLMDGKNKARTVAATNNNETSSRSHAVFTLTQKWH 240
Db 171 HPILGTYVEDLAKLVVRSFOEINLMDGKNKARTVAATNNNETSSRSHAVFTVFTQK 230
Qy 241 DEETKMDTEKVAKISLDVLAGSERATSTGATGARLKEGAEINRSLSLGRVTAALADMS 300
Db 231 DPETNLSTEKVSKISLDVLAGSERADSGAGTRLKEGANINKSLTLTKGVISALARVSK 290
Qy 301 GKQKKNQLVPPYRVSVLTLWLLKDSIGNSMTAMTAISPADINFEETLSLTRYADSK 357
Db 291 -KKKSDFTPYRDSVLTLWLLKENLGNSRTAMTAALSPADINVEETLSLTRYADRAK 346
RESULT 3
Q9ROB4 PRELIMINARY; PRT: 1816 AA.
AC Q9ROB4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE KIFIB.
DE KIFIB.
GN KIFIB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Gong T.L., Winnicki R.S., Kohman D.C., Lomax M.I.;
RT "A novel kinesin of the UNC-104/KIF1 subfamily encoded by the kifb
RL Gene 239:117-127(1999).
DR EMBL; AF090190; AAF06718.1; -.
DR HSRP; P33176; B62.
DR MGD; MGI:108426; Kiflb.
DR INTERPRO: IPR000253; -.
DR INTERPRO: IPR001752; -.
DR INTERPRO: IPR001849; -.
DR PFAM; PF00169; PH; 1.
DR PFAM; PF00225; Kinesin; 1.
DR PFAM; PF00498; FHA; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW SEQUENCE 1816 AA; 204079 MW; E316C295138E5DE CRC64;

```

Query Match 60.5%; Score 1110; DB 11; Length 1816;  
 Best Local Similarity 60.5%; Pred. No. 5.1e-73;  
 Matches 219; Conservative 59; Mismatches 64; Indels 20; Gaps 5;

```

Qy 4 GGNKVVVVRVPFNAREIDRGAKCIVRMGQNTILTPPGAEEKARKSKTMDGPKAFA 63
Db 3 GASVKVAVRVPFNARETSQDAKVVVSQGNQNTTSIINPKE-----AKPSFS 51
Qy 64 FDRSYNSFDKNA-PNYARQEDLPDGLVPLLDNAFKYNNCIFAYGOTSGKSYSMGVG 122
Db 52 FDSYNSHTSTEDPQAFSQQVYRDIQGEMLLHAFEGYNYVCIFAYGOTGAGKSYTMGRQ 111
Qy 123 K--EHWGIPRICQDMFRINELQKDNCTVTVSVYLEIYNERNVRLDLPNPKGNLKVRE 180
Db 112 EPGQOQIVPQCELFEXIND-NCNEEASYSVEVSYMEIYCERVDLLNPKGNLKVRE 170
Qy 181 HPSTGTYVEDLAKLVVRSFOEINLMDGKNKARTVAATNNNETSSRSHAVFTLTQKWH 240
Db 171 HPILGTYVEDLAKLVVRSFOEINLMDGKNKARTVAATNNNETSSRSHAVFTVFTQK 230

```



OS	<i>Caenorhabditis elegans</i> .
OC	Eukaryota; Metazoa; Nematoda; Chromadorae; Rhabditida; Rhabditoidea.
OX	Rhabditidae; Telodoridae; Caenorhabditis.
OC	NCBI_TaxID=6239.
RI	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=94150718; PubMed=7906398;
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA	Crafton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA	Garland A., Grew J., Karsenti J., Kistner J., Kistner N., Latreille J.,
RA	McKern J., Kozlov J., Kozlov J., Kistner J., Kistner N., Latreille J.,
RA	Lightning J., Lloyd C., McMurry A., Mortimore B., O'Callaghan M.,
RA	Parsons J., Percy C., Rifken L., Roopre A., Saunders D., Showkeen R.,
RA	Smaldon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA	Therby-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA	Watson A., Weinstock L., Wilkinson-Spoat J., Wohldman P.;
RT	* 2.2 Mb of contiguous nucleotide sequence from chromosome III of <i>C.</i>
RT	<i>elegans</i> .;
RL	Nature 368:32-38(1994).
RL	(2)
RL	SEQUENCE FROM N.A.

Query Match 60.3%; Score 1106.5; DB 5; Length 1584;  
Best Local Similarity 61.2%; Pred. No. 7.6e-73;  
Matches 218; Conservative 60; Mismatches 61; Indels 17; Gaps 4;

Query Match	60.3%; Score 1106.5; DB 5; Length 1584;
Best Local Similarity	61.4%; Pred. No. 7.6e-7; DB 5; Length 1584;
Matches 218; Conservative	60; Mismatches 61; Indels 17; Gaps 0
OY	6 NIKVWVRPNAREIDRGAKCIVRMGNOTILITPPQCAEKARKSKVITDQGFAPAFD 65
DB	6 NIKVWVRPNAREIDRGAKCIVRMGNOTILITPPQCAEKARKSKVITDQGFAPAFD 65
DB	3 SYKVAVRVRPNOREIENSWKSCVLVGNKTTII-----NCHSINENPSNFD 50
OY	66 RSYWSFKPNAPVABQEDFDLQGLPDNAPKGVNVCNPAYGTCGSKSYVMWYCYK-- 123
DB	66 RSYWSFKPNAPVABQEDFDLQGLPDNAPKGVNVCNPAYGTCGSKSYVMWYCYK-- 123
DB	51 HSYWSFARNDFPIQOVYBELAVEMLEHAFEGNYCIPAYGTCGSKSYVMWYCYK-- 110
OY	124 -EHGVIPIRCODMFRINELQDKNKLNTITVENVSELYEINKEVRDLINPSTGNLKVREHP 182
DB	124 -EHGVIPIRCODMFRINELQDKNKLNTITVENVSELYEINKEVRDLINPSTGNLKVREHP 182
DB	111 DEMGIIPIRLCNDFARION--NNDKDVQVSEVSYMEICYERVDKLLNPSSGNLKVREHP 169
OY	183 STGPIVEDLAKLVFASFOEINLMDGKNKARTVAATNNHSTSSRSIAHFVTLTLTKWIDE 242
DB	183 STGPIVEDLAKLVFASFOEINLMDGKNKARTVAATNNHSTSSRSIAHFVTLTLTKWIDE 242
DB	170 LUGPIYDDITKMAVCVSHDICNLMDEGKNKARTVAATNNHSTSSRSIAHFVTLTKRICA 229
OY	243 ETKNMDEYKAKISLVLDLAGESEARTSTGATGARKEGAENRSJLSTLGRVLAALDMSSGK 302
DB	243 ETKNMDEYKAKISLVLDLAGESEARTSTGATGARKEGAENRSJLSTLGRVLAALDMSSGK 302
DB	230 DSNMLTEKHSKISLVLDLAGESEARTSGAQRKEGANIKSUTTLUGLVTSLKLAESTGK 289
OY	303 OKKNO-LVPIPRDSVUTLLKDSLGSGNSFTAMIAALSPADINPBTLSLTLYADSAK 357
DB	303 OKKNO-LVPIPRDSVUTLLKDSLGSGNSFTAMIAALSPADINPBTLSLTLYADSAK 357
DB	290 KSNKGVPYIPRDSVUTLLRBNAGNSYTNALJANLSPADINFDSTLSLTLYADSAK 345
RESULT	6
RESULT	6

RESULT  
09V7T6  
6

```

ID Q9V7T6 PRELIMINARY: PRT; 1773 AA.
AC Q9V7T6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CG8566 PROTEIN.
GN CG8566.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.G., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jallil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-T., Wasserman D.A., Weisskock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.N., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT Sctence 287, 3185-3195 (1990).
DE Drosophila melanogaster.
DR EMBL: AF003805.3 AKR37937.1; -.
DR HSSP: p17119.3AKR.
DR FLYBASE: FB20034155; CG8566.
DR PROSITE: PS000213; -.
DR INTERPRO: IPR001732; -.
DR INTERPRO: IPR001849; -.
DR PFAM: PF00169; PH; 1.
DR PFAM: PF00225; Kinesin; 1.
DR PFAM: PF00498; FHA; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR SEQUENCE 1773 AA; 200755 MW; 335BE9CD5E4DB681 CRC64;

```

Query Match 60.3%; Score 1106.5; DB 5; Length 1773;  
 Best Local Similarity 62.4%; Pred. No. 8.9e-73;  
 Matches 222; Conservative 52; Mismatches 69; Indels 13; Gaps 4;

```

OY 6 NIKVYVRVPENAEIDRGACIVRMENOTILTPPGAEEKARKSGKTIIMDGPKAFAD 65
DB 32 SVKVAVVRVPENAEIDRGACIVRMENOTILTPPGAEEKARKSGKTIIMDGPKAFAD 83
OY 66 RSYMSFDKNAPYARQEDPOLGLGVPLDPAFGYNNCFAYGOTGSGKSYSMGKGE- 124
DB 84 YSWSHDHDDADESTQSMVYKDIGEMLQHSQDGVNCFAYGOTGAGKSYTMGQREQ 143
OY 125 -HGVIPTQDMFRINELQDKANIKCTVSVYLEINERVRDLINFTKGLKVRHPS 183
DB 144 OEGIIIPICKDLFRITQDTETD-DLKYSVEVSVMYIYCHVRVRLDLPKNGLVREHPL 202
OY 184 TGPVVEDLAKLVRSFOEINLMDGNNKARTVAATNNMSTSSRSRSHAVFTLTLOKHIDEE 243
DB 203 LGPVVEDLSKLVTDYQDIHDLDEGNNKARTVAATNNMSTSSRSRSHAVFTLTORRDLDM 262
OY 244 TKMDTEKVAKISLVLDLAGSERATSGATGARKLKEGAEINRSITLGRVTAALADMSSGKO 303
DB 263 TNLTEKVSKISLVLDLAGSERADSTGAKTRLKEGANINKSLTTLGLVTSALASVASKKK 322
OY 304 --KKNNLPVYRDSVLTLLKDSIGNSMTAMTAASPADINFEETLSTLRVADSAAK 357
DB 323 NTKKADIPYRDSALTYLLRLNGLGNSKTAMTAASPADINFEETLSTLRVADRAK 378
RESULT 7
OYQ9V7E5 PRELIMINARY: PRT; 1816 AA.
AC Q9V7E5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE KIF1B MAJOR ISOFORM.
GN KIF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=BRAIN;
RA Conforti L., Buckmaster A., Tarlton A., Brown M.C., Lyon M.F.,
RA Perry V.H., Coleman M.P.;
RA "The major brain isoform of kif1b lacks the putative mitochondria-
RT binding domain";
RL Mamm. Genome 10:0-0(1999).
DR EMBL: AF131865; AAD3438.1; -.
DR HSSP: P33176; 18G2.
DR MGD: MGI:108426; Kif1b.
DR INTERPRO: IPR000253; -.
DR INTERPRO: IPR001752; -.
DR INTERPRO: IPR001849; -.
DR PFAM: PF00169; PH; 1.
DR PFAM: PF00225; Kinesin; 1.
DR PFAM: PF00498; FHA; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR PROSITE: PS50067; FHA_DOMAIN; 1.
DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DE Kif1b protein; Microtubules; ATP-binding; Coiled coil.
KW SEQUENCE 1816 AA; 204153 MW; 93CE1969DB4895A CRC64;

```

Query Match 60.1%; Score 1102; DB 11; Length 1816;  
 Best Local Similarity 60.2%; Pred. No. 2e-72;  
 Matches 218; Conservative 59; Mismatches 65; Indels 20; Gaps 5;

OY 4 GSNKYVVRVPENAEIDRGACIVRMENOTILTPPGAEEKARKSGKTIIMDGPKAF 63  
 DB 3 GASVAVVRVPENAEIDRGACIVRMENOTILTPPGAEEKARKSGKTIIMDGPKAF 51

Qy	241	DRETMNDTEKVKAKISLVDLGASERATSGTARLKGAEINRSLTIGRTVLAADKSS	300
Db	230	DOLTGDSLSEKVSISLVNLGASERADSSGARGNRLKSGANINRSLTIGKVTSLADLQS	289
Qy	301	GQKKNGLVPRSVLSTMLKDSLGNSWMTAAIISPADINFEETLSTLRYADSAK	357
Db	290	-KKRKSDFTPYRDSVLTMLKENIGNSRTAAIISPADINFEETLSTLRYADTK	345
RESULT	9		
QNGQ02		PRELIMINARY; PRT; 2205 AA.	
ID	QNGQ02		
DC	QNGQ02		
DT	01-OCT-2000 (TREMBL-rel. 15, Created)		
DT	01-OCT-2000 (TREMBL-rel. 15, Last sequence update)		
DT	01-OCT-2000 (TREMBL-rel. 15, Last annotation update)		
DT	KINETSIN UNC104/KIFIA HOMOLOG.		
DE	UNC104.		
GN	Dictyostelium discoideum (Slime mold).		
OS	Eukaryota; Dictyostelida; Dictyostelium.		
OX	NCBI_Taxid:44689;		
RB	SEQUENCE FROM N.A.		
EX	WEDLIN-2001:4990; PubMed-10545495;		
RA	Pollock N., de Hostos E.L., Turck C.W., Vale R.D.,		
RT	*Reconstitution of membrane transport powered by a novel dimeric		
RT	kinesin motor of the Unc104/KIFIA family purified from		
RT	Dictyostelium.*;		
RT	J. Cell Biol. 147:493-506(1999).		
RP	[2]		
RN	SEQUENCE FROM N.A.		
RL	Pollock N., Vale R.D.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RD	EMBL: AF245247.1; AAF65384.1;		
SD	SEQUENCE 2205 AA; 248001 MW; 02C5101E9D61C94D CRC64;		
Query Match	58.9%; Score 1081; DB 5; Length 2205;		
Best Local Similarity	62.1%; Pred. No. 9.2e-71;		
Matches	221; Conservative 49; Mismatches 78; Indels 8; Gaps		
Qy	6	NKTVYVVRPNAREIDRGACIVRMENQOTILTPPGAEKARKSGKTIIMPCKAFAD	65
Db	2	NVQAVVRPNRSEKKNALIVQNNKSTILITPSAL--RANLAPTPADDEKSEFSD	59
Qy	66	RSYSGFKAWYARQEDLPDQLGVLDNAFGVYNCTIAYGQTGSGKSYSMWGKHEH	125
Db	60	YSYVSDISDHPFAQSOSTVYNDLQKEVLANRWDGFCISFAYQVTSGRSISMGIGEK	119
Qy	126	GVPIPCIDQDFRINELQDKN--LCTVRYVSYTIFINRVRVLDNPSTK--GNLVRKH	181

Qy	182	PSTGQYVEDJAKLVYRSPQIENLADGQKARVAAFNANWETSSASHAVPTTUTQKWD	241
Db	183		242
Db	184		243
Db	180	PSTGQYVEDJAKLVYRSPQIENLADGQKARVAAFNANWETSSASHAVPTTUTQKWD	239
Qy	242	BETKMDTBRVAKLSVLVDAGSERATSTGATGARKLKEGAEINRSLTTLGRVYAAALADKSSG	301
Db	240	KTRGTAIDRVSKLSVLVDAGSRANSTGATGVRLKEGANIKSLTTLGKYVIALAENST	298
Qy	302	KQKKNQVLVYRDDSVYTLMLKDSLOGNSMTAMTAATSPADINFEPTLSTLRYADSAK	357
Db	299	-SKAAVFPYRDSVYTLMLKDSLOGNSMTAMTAATSPADINFEPTLSTLRYADSAK	353
RESULT	10		
AC	08658	PRELIMINARY;	PRT: 689 AA.
ID	08658		
IC	08658		
DT	01-NOV-1998 (TREMblrel. 08. Created)		
DT	01-NOV-1998 (TREMblrel. 08. Last sequence update)		

01-OCT-2000 (TREMBLER)	15	Lat annotation update)
DT	DE	KINESIN-LIKE PROTEIN KIF1B (FRAGMENT).
GN	GN	KIF1B
OS	OS	Rattus norvegicus (Rat)
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
OX	OX	NCBI_TaxID=10116;
LN	LN	[1]
RP	RP	SEQUENCE FROM N.A.
RA	RA	Faire K., Gruber D., Bulinski J.C.;
RT	RT	"Identification of kinesin-like molecules in myogenic cells.";
RL	RL	Eur. J. Cell Biol. 0-0(1998).
DR	DR	EMBL; AF083331; AAC33292.1; -.
DR	DR	HSSP; P31176; 1RG2
DR	DR	INTERPRO; IPR00253; -.
DR	DR	INTERPRO; IPR001356; -.
DR	DR	INTERPRO; IPR001752; -.
DR	DR	PFAM; PF00225; kinesin; 1.
DR	DR	PFAM; PF00498; FHA; 1.
DR	DR	PRINTS; PR00380; KINESINHEAVY.
DR	DR	PROSITE; PS00027; HOMEBOX_1; UNKNOWN_1.
DR	DR	PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
DR	DR	NON_TER 689 689
FT	FT	SEQUENCE 689 AA; 777352 MW; F64818C12FE69319 CRG64;

Query Match 58.5%; Score 1073.5; DB 11; Length 689;  
Best Local Similarity 60.2%; Pred. No. 6.1e-71;  
Matches 216; Conservative 57; Mismatches 67; Indels 19; Gaps 7;

QY	63	4	GN1KVVVVVRFPFNARIDRGAACKIVMBRGNQTLTPPPGAEBKARKSGKTIIDGPKAPA	63
Db	64	3	GASVAVAVRVPFNSRSTKESKCIQIOMGNSTSIINPNPKE	51
QY	122	64	FRDSYWSF-DKNAFNWAROEDFDJGLVPLLDNFAFYGNCCIFAYGOTGSGKSYSMWGY	122
Db	123	52	FQYSYWSHSPDPCFASQSRVYNDIKEMLLHAFEGYNVICIFAYQOTGASGVYTMWQK	111
QY	180	123	KEH--GVTRICOMFERRNELQKDLNACTIVEYSLEYTVNRRVOLLNPTSTKGNLKVRE	180
Db	181	112	ESQNGHIIPTCEBSELPKND-NCNEDMSYSV-SYMEICYEKVROLLNPKNKNLRYRE	169
QY	240	181	HPSTGPPVYEDLAKLVRSQETELNMDQENKARVVAATNNHSTSSSHAVPTLTFLQKH	240
Db	241	170	DEFTKMDTEKVAKLSLVLDLGRASRGATGATGABLVKGEENINELSTLGRVYIAALDQSS	239
QY	300	241	HPSTGPPVYEDLAKLVRSQETELNMDQENKARVVAATNNHSTSSSHAVPTLTFLQKH	300
Db	289	230	DPETNLSTKVTKLSLVLDLGRASRGATGATGABLVKGEENINELSTLGRVYIAALDQSS	289
QY	357	301	GKOKNOLVPPYRDSVLTWLLKXSLGNGSMTAMIAALSPADINPEETL--STLRYADSAK	357
Db	347	290	-KKKTFDTPYRDSVLTWLLRENGJNSMTAWAALSADINPYDETSLSTLRYADRAK	347

RESULT	11
801349	
IC	O01349 PRELIMINARY; FRT; 1921 AA.
AC	O01349;
DT	01-JUL-1997 (TReMBLrel_04, Created)
DT	01-JUL-1997 (TReMBLrel_04; Last sequence update)
DE	01-OCT-2000 (TReMBLrel_15, Last annotation update)
DE	KNESIN-73
DS	Drosophila melanogaster (Fruit fly).
OS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephyraoidea; Drosophilidae; Drosophila.
RN	NCBI_TaxID=7227;
OX	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=97188425; PubMed=9037010;
RY	Li H.P., Liu Z.M., Nirenberg M.;
RT	"kinesin-73 in the nervous system of Drosophila embryos.";

PROC. Natl. Acad. Sci. U.S.A. 94:1086-1091(1997).

RL  
DR ESBL1: 03178; KAG5404.1; -;  
DR ESBL2: 03179; KAG5404.1; -;  
DR ESBL3: 03180; KAG5404.1; -;  
DR FLYBASE: FB0010958; Klc-73.  
DR IPRO00338; -;  
DR IPRO00339; -;  
DR IPRO01220; -;  
DR INTERPRO; -;  
DR INTERPRO; IPRO01752; -;  
DR IPAM; pf00225; kinesin; 1.  
DR PFAM; pf01302; CAP GLY; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR PROSITE; PS00307; LECTIN LEGUME\_BETA; UNKNOWN\_1.  
DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN; 1.  
DR PROSITE; PS00845; CAP GLY; 1.  
DR PROSITE; PS00667; KINESIN\_MOTOR\_DOMAIN; 1.  
DR Motor protein; Microtubules; AMP-binding; Coiled coil.  
SQ SEQUENCE 1931 AA; 215047 MW; 4643F6F9783E99D0 CRC64.

Query Match	56.1%;	Score 1028;	DB 5;	Length 1921;
Best Local Similarity	60.3%;	Pred. No. 6.1e-67;		
Matches 213; Conservative	45;	Mismatches 85;	Indels 10;	Gaps 4;

[illegible]

RESULT	12
Q9W7C9	
ID	Q9W7C9;
AC	Q9W7C9;
DT	01-MAY-2000 (tREMBrel. 13, Created)
DT	01-MAY-2000 (tREMBrel. 13, Last sequence update)
DT	01-OCT-2000 (tREMBrel. 15, Last annotation update)
GN	KRC-73 PROTEIN.
GN	KRC-73

OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Eohydroidae; Prosophilidae; Prosophila.

11	SEQUENCE FROM N. A.	
12	RC	STRAINS-BERKLEY.
13	RC	
14	RC	
15	RC	
16	RC	
17	RC	
18	RC	
19	RC	
20	RC	
21	RC	
22	RC	
23	RC	
24	RC	
25	RC	
26	RC	
27	RC	
28	RC	
29	RC	
30	RC	
31	RC	
32	RC	
33	RC	
34	RC	
35	RC	
36	RC	
37	RC	
38	RC	
39	RC	
40	RC	
41	RC	
42	RC	
43	RC	
44	RC	
45	RC	
46	RC	
47	RC	
48	RC	
49	RC	
50	RC	
51	RC	
52	RC	
53	RC	
54	RC	
55	RC	
56	RC	
57	RC	
58	RC	
59	RC	
60	RC	
61	RC	
62	RC	
63	RC	
64	RC	
65	RC	
66	RC	
67	RC	
68	RC	
69	RC	
70	RC	
71	RC	
72	RC	
73	RC	
74	RC	
75	RC	
76	RC	
77	RC	
78	RC	
79	RC	
80	RC	
81	RC	
82	RC	
83	RC	
84	RC	
85	RC	
86	RC	
87	RC	
88	RC	
89	RC	
90	RC	
91	RC	
92	RC	
93	RC	
94	RC	
95	RC	
96	RC	
97	RC	
98	RC	
99	RC	
100	RC	

[illegible]

Db	234	LIDAKSTSGEKVKGLSDWLAGSRATKTAAGDRKLRKSGNSTTUGLVISLADQ	293
Qy	299	SSGCKKNQVLPVRSVLTMLLKSLGGNSKMTIAASTSPADNFEETLSLTGYADSAK	357
Db	294	SAGK-NKNEFVPRSVLTMLLKSLGGNSKMTAVATSPADNYDETLSLTGYADRAK	351
	RESULT	14	
	Q20888		
	ID	Q20888	PRT: 1576 AA.
	AC	Q20888:	PRELIMINARY;
	DT	01-NOV-1996	(TrEMBLrel. 01, Created)
	DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)
	DT	01-JUN-2000	(TrEMBLrel. 01, Last annotation update)
	DT	01-JUN-2000	(TrEMBLrel. 01, Last sequence update)
	GN	F5663.3	SIMILAR TO KINESIN-RELATED PROTEIN. NCBI GI: 1109842.
	GN	F5663.3	Caenorhabditis elegans.
	OC	Karyophilids; Metazoa; Chromadorea; Rhabditida; Rhabditidae;	
	OC	Caenorhabditis; Rhabditidae; Rhabditinae; Caenorhabditis.	
	OC	NCBI TaxID=6239,	
	PN	[1]	SEQUENCE FROM N.A.
	RP	STRAIN=BRISTOL N2.	
	RX	WILSON R., Alnsworth K., Anderson K., Baynes C., Berks M.,	
	RX	Medline=94150718; PubMed=7906398;	
	RA	Bonfield J., Burton J., Connell M., Copsey T., Coulson A.,	
	RA	Croxson M., Dear S., Du Z., Durbin R., Favello A., Fulton L.	

Db	234	LIDAKSTSGEKVKGLSDWLAGSRATKTAAGDRKLRKSGNSTTUGLVISLADQ	293
Qy	299	SSGCKKNQVLPVRSVLTMLLKSLGGSNKMTIAATSPADNFEETLSLTGYADSAK	357
Db	294	SAGK-NKNEFVPRSVLTMLLKSLGGSNKMTAVATSPADNYDETLSLTGYADRAK	351
	RESULT	14	
	Q20888		
	ID	Q20888	PRT: 1576 AA.
	AC	Q20888:	PRELIMINARY;
	DT	01-NOV-1996	(TRENBLREL. 01, Created)
	DT	01-NOV-1996	(TRENBLREL. 01, Last sequence update)
	DT	01-JUN-2000	(TRENBLREL. 01, Last annotation update)
	DT	01-JUN-2000	(TRENBLREL. 01, Last sequence update)
	GN	F5663.3	SIMILAR TO KINESIN-RELATED PROTEIN. NCBI GI: 1109842.
	GN	F5663.3	Caenorhabditis elegans.
	OC	Karyophilids; Metazoa; Chromadorea; Rhabditida; Rhabditidae;	
	OC	Caenorhabditis; Rhabditidae; Rhabditinae; Caenorhabditis.	
	OC	NCBI TaxID=6239,	
	PN	[1]	SEQUENCE FROM N.A.
	RP	CP	STRAIN-BRISTOL N2.
	RA	RA	WILSON R., Alnsworth K., Anderson K., Baynes C., Berks M.,
	RA	RA	Benfield J., Burton J., Connell M., Copsey T., Cooper A.,
	RA	RA	Croxson M., Dear S., Du Z., Durbin R., Favello A., Fulton L.

Query Match	52.6%	Score	965.5	DB	5	Length	1576
Best Local Similarity	53.9%	Pred.	No. 1.8e-82				
Matches	208	Conservative	45	Mismatches	82	Indels	51
Gaps							
QY	7	IKWVVRPNAREIDRCACIVIRMEGNQIILPPPPGAEEKARKSGKTIWDGPKFAFDR	66				
DB	10	KVALIRVRPNKRELDUKTSVVRIOEQCVLHHP--LEEK-----NSKTFITFDH	57				
QY	67	SWTFSDKNAPYARQEDLFODGLVPLLDNAFKYNNCIFAYGOTGSKYSMMGYKEHG	136				
DB	58	SPCSTDPHSYDFASQETSVYHLGSGVGNFAPGYNACIFAYGOTGSKYSMMGTPDQPG	117				
QY	127	VPIRCQMPRRINELQKDKNLCTVSVSYLEYIYERNVROLLNPSTKGN-LKYREHPSTG	185				
DB	118	IIPRVCMNDIFTRIQQ--TNSNSQFVKEVSYMEYIYERNVROLLDPKSKKALVREHKILG	176				
QY	186	PVYEDLAKLVRSFOETENIMDEGNKARTVAATNNNETSSRSRAVFTLTITQKWDHEETK	245				
DB	177	PWYDGLSTIAVNSFQISNLLEGNKASTVAATNNNAESSRSRAVFSLTIVTQTLHDLNG	236				
QY	246	MOTEKVAKISLVLDLAGSERATSTGATGARLUKEAGIN-----	282				
DB	237	FSGEKVAKISLVLDLAGSERAGTGAVKRLEEGGNINKNLVSLFRNLEKIKDFKPSAD	296				
QY	283	-----RSLSTLGRVIAALADSSGQKKNQLPYPRDSVLITWLLKDSIGNSMTA	331				
DB	297	VYVVFYNOKLDFRSLTGMVISAIAERNS-----KKDKPIPYRDSVLITWLLKDSIGNSRTV	353				
QY	332	MIAATSPADINFEETLSTLRVADSAAK	357				
DB	354	MIATSPAADNYEETLSTLRVADRAAK	379				

Search completed: April 25, 2001, 10:18:12  
Job time: 318 sec

Search completed: April 25, 2001, 10:18:12  
Job time: 318 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:12:47 ; Search time 56.07 Seconds  
(without alignments)  
186,568 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_602\_784

Perfect score: 938

Sequence: 1 QEOSLLRHSVNSQLSGSPAP.....ELRQQQAWEALKTAQSF 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 segs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Genseq\_0401.\*

- 1: /SIDSL1/gcgdata/geneseq/AA1980.DAT.\*
- 2: /SIDSL1/gcgdata/geneseq/AA1981.DAT.\*
- 3: /SIDSL1/gcgdata/geneseq/AA1982.DAT.\*
- 4: /SIDSL1/gcgdata/geneseq/AA1983.DAT.\*
- 5: /SIDSL1/gcgdata/geneseq/AA1984.DAT.\*
- 6: /SIDSL1/gcgdata/geneseq/AA1985.DAT.\*
- 7: /SIDSL1/gcgdata/geneseq/AA1986.DAT.\*
- 8: /SIDSL1/gcgdata/geneseq/AA1987.DAT.\*
- 9: /SIDSL1/gcgdata/geneseq/AA1988.DAT.\*
- 10: /SIDSL1/gcgdata/geneseq/AA1989.DAT.\*
- 11: /SIDSL1/gcgdata/geneseq/AA1990.DAT.\*
- 12: /SIDSL1/gcgdata/geneseq/AA1991.DAT.\*
- 13: /SIDSL1/gcgdata/geneseq/AA1992.DAT.\*
- 14: /SIDSL1/gcgdata/geneseq/AA1993.DAT.\*
- 15: /SIDSL1/gcgdata/geneseq/AA1994.DAT.\*
- 16: /SIDSL1/gcgdata/geneseq/AA1995.DAT.\*
- 17: /SIDSL1/gcgdata/geneseq/AA1996.DAT.\*
- 18: /SIDSL1/gcgdata/geneseq/AA1997.DAT.\*
- 19: /SIDSL1/gcgdata/geneseq/AA1998.DAT.\*
- 20: /SIDSL1/gcgdata/geneseq/AA1999.DAT.\*
- 21: /SIDSL1/gcgdata/geneseq/AA2000.DAT.\*
- 22: /SIDSL1/gcgdata/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	938	100.0	784	20 Y06618	Thermomyces lanuginosus
2	102.5	10.9	436	21 G43544	Arabidopsis thaliana
3	92.5	9.9	264	21 B43550	Human cancer assoc
4	91	9.7	592	21 G15484	Arabidopsis thaliana
5	91	9.7	632	21 G46578	Arabidopsis thaliana
6	91	9.7	644	21 G46577	Arabidopsis thaliana
7	91	9.7	650	21 G15483	Arabidopsis thaliana
8	91	9.7	702	21 G46576	Arabidopsis thaliana
9	91	9.7	725	21 G15482	Arabidopsis thaliana
10	90.5	9.6	326	21 G35664	Arabidopsis thaliana
11	90.5	9.6	1816	21 Y95440	Caenorhabditis eleg

12	89.5	9.5	590	20 Y41101	Rat gadd34 protein
13	89.5	9.5	1028	14 R41044	Plasmomium vivax D
14	89.5	9.5	1115	12 R13457	Duffy antigen
15	89.5	9.5	1115	18 W22478	P. vivax DABP bind
16	89.5	9.5	1115	21 Y77899	TNF-R-PI. vivax Du
17	89.5	9.5	1245	16 R70106	Rat HNG2 (RUMK1) p
18	88	9.4	404	21 B19686	Rat HNG2 (RUMK3) p
19	88	9.4	427	21 B19687	Rat HNG2 (RUMK1) p
20	88	9.4	665	21 B19685	Staphylococcus aur
21	87.5	9.3	793	21 B38455	Human Fickart12 pr
22	87	9.3	936	16 W89801	Human Fickart12 pr
23	87	9.3	936	16 W89801	Arabidopsis thaliana
24	86.5	9.2	302	21 G35665	Arabidopsis thaliana
25	86.5	9.2	340	19 W59952	Amino acid sequenc
26	86.5	9.2	623	21 G42595	Arabidopsis thaliana
27	86.5	9.2	635	21 G42594	Arabidopsis thaliana
28	86.5	9.2	693	21 G42593	Arabidopsis thaliana
29	85.5	9.1	2100	20 W89579	Calcium permeable
30	84	9.0	653	21 Y94907	Human secreted pro
31	84	9.0	659	19 W37724	CD2 associated int
32	84	9.0	665	21 B19690	Human HNG2 (short
33	84	9.0	709	21 B19691	Human HNG2 (long f
34	83	8.8	1724	21 G31229	Arabidopsis thaliana
35	83	8.8	1730	21 G31228	Arabidopsis thaliana
36	83	8.8	1739	21 G50122	Arabidopsis thaliana
37	83	8.8	1745	21 G50121	Arabidopsis thaliana
38	83	8.8	1777	21 G31220	Arabidopsis thaliana
39	82.5	8.8	257	21 G35740	Arabidopsis thaliana
40	82.5	8.8	258	21 G35779	Arabidopsis thaliana
41	82.5	8.8	1056	21 G31894	Arabidopsis thaliana
42	82.5	8.8	1073	21 G31893	Arabidopsis thaliana
43	82.5	8.8	1079	21 G48607	Arabidopsis thaliana
44	82.5	8.8	1087	21 G31892	Arabidopsis thaliana
45	82.5	8.8			

# ALIGNMENTS

RESULT	1
Y06618	ID Y06618 standard; Protein; 784 AA.
XX	XX Y06618;
XX	XX Y06618;
XX	XX Y06618;
DT	26-OCT-1999 (first entry)
XX	XX Thermomyces lanuginosus kinesin motor protein TL-gamma.
DE	XX
XX	XX TL-gamma; kinesin; motor protein; microtubule; unc-104; Infection;
XX	XX neurodegenerative disease; Alzheimer's disease;
KW	KW Parkinson's disease; Huntington's disease;
KW	KW amyotrophic lateral sclerosis.
XX	XX Thermomyces lanuginosus.
OS	XX
XX	XX W09937659-A1.
PN	XX
PD	XX 29-JUL-1999.
XX	XX
XX	XX 22-JAN-1999; 99WO-US01355.
XX	XX
PR	23-JAN-1998; 98US-0072361.
XX	XX (RECC ) UNIV CALIFORNIA.
PA	XX
XX	XX Goldstein LSB, Sakowicz R;
PI	XX
XX	XX WPI; 1999-493950/41.
XX	XX N-PSDB: X87656.
DR	XX
XX	XX New nucleic acid encoding microtubule motor protein, used for
XX	XX diagnosis of fungal infection and neurodegenerative disease
PT	XX

```
XX Claim 5; Page 70-71: 75pp: English.
XX PS
XX CC This sequence represents Thermomyces lanuginosus TL-gamma, a novel
XX CC ATP-dependent, plus end-directed microtubule motor protein that is
XX CC a member of the unc-104 family and kinesin superfamily. The
XX CC invention provides TL-gamma nucleic acids (see X87656), proteins
XX CC and antibodies, and methods of screening for TL-gamma modulators
XX CC potentially useful for treating hyphal fungal infections and
XX CC diseases caused by mutated TL-gamma, e.g. neurodegeneration
XX CC involving anterograde axonal transport, such as Alzheimer's,
XX CC Parkinson's or Huntington's diseases or amyotrophic lateral
XX CC sclerosis. Detection of TL-gamma allows differentiation between
XX CC hyphal and non-hyphal fungal infections.
XX Sequence 784 AA:
SQ
Query Match 100.0%; Score 938; DB 20; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.6e-86;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QEQSLRHVSNTSQLGSPAFGRHRTLSKAGSDADGDSKSDSPLPFRGKDSDFYARRE 60
DB 602 gqgslrhvntsqlgspagprhrtlskagdadgdsrdsplphfrgkdsdfyarre 661
QY 61 AASALIGLDQKISHLTDELDALFDYOKARVRGLVEDNEDSDSQSFPVRDKTMSNG 120
DB 662 aasailgldqkishltddelaldfdvqkaravrrglvedneddsqgsfpvrckymnsng 721
QY 121 TIDNFSLOTATMPGTPRSSDDGDALFFGDKSKSDASNVVDVEELRQQQAQMEELKTKA 180
DB 722 tidnfsldtaltmpdprsdgdgdalffgdksksdasnvvdeelrqggaqmeelktak 781
QY 181 QEF 183
DB 782 qef 784
RESULT 2
G43544
XX ID G43544 standard; Protein; 436 AA.
XX AC G43544;
XX DE 18-OCT-2000 (first entry)
XX KW Arabidopsis thaliana protein fragment SEQ ID NO: 54439.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridization assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PS 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142603.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
```

PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	25-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match			10.9%; Score 102.5; DB 21; Length 436;
Best Local Similarity			24.6%; Pred. No. 0.035; 53; Indels 7; Gaps
Matches 28; Conservative 26; Mismatches 37;			
QY	69	DOIKSHLTDLDALFDVOKARVRGLVEDNEDSDSSFPVRKYMSKGTIDNFSLD 128	
Dy	35	dqpdgdtdevdsmddgeseeddqgdtceddgdsceedgenkedgedesdfe-- 92	
QY	129	TATWTFGRPSDDGDLFLFGDKSKQDASNVDELRQQAQMEALKTAQOE 182	
Dy	93	-----dgnkdesegdgndnkdagmeelekevksrqgdlnkkrdkge 141	
RESULT	3		
B43650			
ID	B43650 standard; Protein; 264 AA.		
AX	B43650;		
DT	08-FEB-2001 (first entry)		
DE	Human cancer associated protein sequence SEQ ID NO:1095.		
XX	Human; cancer associated gene; cancer antigen; detection; cancer;		
XX	diagnosis; cytostatic; proliferative; vulnary; immunomodulator;		
KW	antidiabetic; antischismatic; antirheumatic; antithrombotic; antiviral;		
KW	antiinflammatory; antithyroid; antiallergic; antibacterial; cardiatic;		
KW	dermatological; neuroprotective; thrombolytic; coagulant; neotropic;		
KW	vasotonic; antipsoriatic; antiangiogenic; gene therapy; inflammation;		
KW	immune disorder; haemacopoietic cell disorder; autoimmune disorder;		
KW	allergic reaction; graft versus host disease; organ rejection;		
KW	hemocytocytic; cardiovascular; neurovascular disorder; infection;		
KW	neurological; disease; drug screening.		
OS	Homo sapiens.		
PN	W02000053530-A1.		
PN	21-SEP-2000.		
PD			
PF	08-MAR-2000; 2000WO-US05882.		
PR			
XX	12-MAR-1999; 99US-0124270.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX	Rosen CA, Ruben SM;		
PI	WPI: 2000-587533/55.		
DR	NF-PUB: C77859.		





PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
XX	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
XX	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542.
PR	99US-01231825.	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	99US-01233180.	PR	16-JUL-1999;	99US-0144085.
PR	09-MAR-1999;	PR	19-JUL-1999;	99US-0144086.
PR	99US-01235788.	PR	19-JUL-1999;	99US-0144325.
PR	23-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	99US-0126264.	PR	19-JUL-1999;	99US-0144332.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	99US-0126785.	PR	19-JUL-1999;	99US-0144334.
PR	01-APR-1999;	PR	20-JUL-1999;	99US-0144335.
PR	99US-0127462.	PR	20-JUL-1999;	99US-0144352.
PR	06-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	99US-0128234.	PR	21-JUL-1999;	99US-0144884.
PR	08-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	16-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	99US-0128714.	PR	21-JUL-1999;	99US-0145088.
PR	19-APR-1999;	PR	22-JUL-1999;	99US-0145085.
PR	99US-0129845.	PR	22-JUL-1999;	99US-0145087.
PR	15-APR-1999;	PR	22-JUL-1999;	99US-0145089.
PR	99US-0130077.	PR	22-JUL-1999;	99US-0145192.
PR	21-APR-1999;	PR	23-JUL-1999;	99US-0145145.
PR	99US-0130449.	PR	23-JUL-1999;	99US-0145218.
PR	23-APR-1999;	PR	23-JUL-1999;	99US-0145224.
PR	99US-0130510.	PR	26-JUL-1999;	99US-0145276.
PR	23-APR-1999;	PR	27-JUL-1999;	99US-0145913.
PR	99US-0130891.	PR	27-JUL-1999;	99US-0145918.
PR	28-APR-1999;	PR	27-JUL-1999;	99US-0145919.
PR	99US-0131449.	PR	28-JUL-1999;	99US-0145951.
PR	30-APR-1999;	PR	02-AUG-1999;	99US-0146386.
PR	99US-0132048.	PR	02-AUG-1999;	99US-0146388.
PR	04-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	99US-0132485.	PR	02-AUG-1999;	99US-0147038.
PR	06-MAY-1999;	PR	03-AUG-1999;	99US-0147204.
PR	99US-0132486.	PR	04-AUG-1999;	99US-0147302.
PR	99US-0132487.	PR	05-AUG-1999;	99US-0147192.
PR	07-MAY-1999;	PR	06-AUG-1999;	99US-0147260.
PR	99US-0132863.	PR	06-AUG-1999;	99US-0147303.
PR	99US-0134256.	PR	06-AUG-1999;	99US-0147416.
PR	14-MAY-1999;	PR	09-AUG-1999;	99US-0147493.
PR	99US-0134218.	PR	09-AUG-1999;	99US-0147935.
PR	99US-0134219.	PR	10-AUG-1999;	99US-0148171.
PR	14-MAY-1999;	PR	10-AUG-1999;	99US-0148319.
PR	99US-0134221.	PR	12-AUG-1999;	99US-0148341.
PR	14-MAY-1999;	PR	12-AUG-1999;	99US-0148565.
PR	99US-0134370.	PR	13-AUG-1999;	99US-0148684.
PR	18-MAY-1999;	PR	16-AUG-1999;	99US-0149368.
PR	99US-0134768.	PR	17-AUG-1999;	99US-0149175.
PR	99US-0134941.	PR	18-AUG-1999;	99US-0149426.
PR	20-MAY-1999;	PR	20-AUG-1999;	99US-0149722.
PR	99US-0135124.	PR	20-AUG-1999;	99US-0149929.
PR	99US-0135353.	PR	23-AUG-1999;	99US-0149902.
PR	24-MAY-1999;	PR	25-AUG-1999;	99US-0149520.
PR	99US-0135629.	PR	26-AUG-1999;	99US-0150586.
PR	25-MAY-1999;	PR	26-AUG-1999;	99US-0151085.
PR	99US-0136021.	PR	27-AUG-1999;	99US-0151086.
PR	27-MAY-1999;	PR	27-AUG-1999;	99US-0151087.
PR	99US-0136392.	PR	30-AUG-1999;	99US-0151088.
PR	99US-0136782.	PR	31-AUG-1999;	99US-0151438.
PR	01-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	99US-0137222.	PR	01-SEP-1999;	99US-0151931.
PR	99US-0137528.	PR	10-SEP-1999;	99US-0152363.
PR	04-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	99US-0137502.	PR	15-SEP-1999;	99US-0154018.
PR	07-JUN-1999;	PR	16-SEP-1999;	99US-0154039.
PR	99US-0137724.			
PR	99US-0138094.			
PR	08-JUN-1999;			
PR	99US-0138540.			
PR	10-JUN-1999;			
PR	99US-0138847.			
PR	11-JUN-1999;			
PR	99US-0139115.			
PR	16-JUN-1999;			
PR	99US-0139422.			
PR	16-JUN-1999;			
PR	99US-0139453.			
PR	17-JUN-1999;			
PR	99US-0139452.			
PR	18-JUN-1999;			
PR	99US-0139455.			
PR	18-JUN-1999;			
PR	99US-0139456.			
PR	18-JUN-1999;			
PR	99US-0139457.			
PR	18-JUN-1999;			
PR	99US-0139458.			
PR	18-JUN-1999;			
PR	99US-0139460.			
PR	18-JUN-1999;			
PR	99US-0139461.			
PR	18-JUN-1999;			
PR	99US-0139462.			
PR	18-JUN-1999;			
PR	99US-0139463.			
PR	18-JUN-1999;			
PR	99US-0139750.			
PR	21-JUN-1999;			
PR	99US-0139817.			
PR	22-JUN-1999;			
PR	99US-0139899.			
PR	23-JUN-1999;			
PR	99US-0140353.			
PR	23-JUN-1999;			
PR	99US-0140354.			
PR	24-JUN-1999;			
PR	99US-0140695.			
PR	28-JUN-1999;			
PR	99US-0140823.			
PR	29-JUN-1999;			
PR	99US-0140991.			
PR	30-JUN-1999;			
PR	99US-0141287.			
PR	01-JUL-1999;			
PR	99US-0141842.			



```

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 08-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148541.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148984.
PR 16-AUG-1999; 99US-0149178.
PR 17-AUG-1999; 99US-0149178.
PR 18-AUG-1999; 99US-0149178.
PR 20-AUG-1999; 99US-0149726.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149820.
PR 23-AUG-1999; 99US-0149820.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150560.
PR 26-AUG-1999; 99US-0150886.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

```

```

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 9.7%; Score 91; DB 21; Length 644;

Best Local Similarity 23.9%; Pred. No. 0.86;

Matches 45; Conservative 30; Mismatches 75; Indels 38; Gaps 9;

```

Qy 21 GRHDRTLSKAGSDADGSDSDSPHFPRKQSDW----FYAREASAAIIGLDQKISHL 75
Db 259 gvhdsksksatpapsprpsppppssgsgslgafwatghaktsvvsednk---- 314
Qy 76 TDELDALFDVQKARAVRGLVEDNEDSDSQSPFVRDKYMSNGTIDNFSLOTATMPG 135
Db 315 nmpkfdepnatskservr---vdshq---pkpbspvrge--arglqknkldettisqkn 366
Qy 136 TP-----RSD-----DGDALFF--GDKKSKODASNVDEELRQO---QOQWEE 174
Db 367 ttpaanmrtrvskddafnsfvadfttdtkfdngnkpqkealeaeigrkldelqtksek 426
Qy 175 ALKTAQBE 182
Db 427 aeitakfe 434

```

RESULT 7

GI5483  
 ID GI5483 standard; Protein: 650 AA.

XX  
 AC  
 GI5483;  
 XX



DT	17-OCT-2000	(first entry)	
XX	Arabidopsis thaliana	protein fragment SEQ ID NO: 15752.	
XX	Protein identification: signal transduction pathway; metabolic pathway;		
XX	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
XX	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000;	2000P-0301439.	
XX	25-FEB-1999;	99US-0121825.	
XX	05-MAR-1999;	99US-0123180.	
XX	09-MAR-1999;	99US-0123548.	
XX	25-MAR-1999;	99US-0125788.	
XX	29-MAR-1999;	99US-0126785.	
XX	01-APR-1999;	99US-0126785.	
XX	06-APR-1999;	99US-0126234.	
XX	08-APR-1999;	99US-0128714.	
XX	16-APR-1999;	99US-0129845.	
XX	19-APR-1999;	99US-0130077.	
XX	21-APR-1999;	99US-0130449.	
XX	23-APR-1999;	99US-0130510.	
XX	28-APR-1999;	99US-0130891.	
XX	30-APR-1999;	99US-0131449.	
XX	30-APR-1999;	99US-0132048.	
XX	04-MAY-1999;	99US-0132407.	
XX	05-MAY-1999;	99US-0132484.	
XX	06-MAY-1999;	99US-0132485.	
XX	07-MAY-1999;	99US-0132486.	
XX	11-MAY-1999;	99US-0132563.	
XX	14-MAY-1999;	99US-0132456.	
XX	14-MAY-1999;	99US-0134218.	
XX	14-MAY-1999;	99US-0134219.	
XX	14-MAY-1999;	99US-0134370.	
XX	18-MAY-1999;	99US-0134768.	
XX	19-MAY-1999;	99US-0134941.	
XX	21-MAY-1999;	99US-0135124.	
XX	24-MAY-1999;	99US-0135353.	
XX	25-MAY-1999;	99US-0135629.	
XX	27-MAY-1999;	99US-0136021.	
XX	28-MAY-1999;	99US-0136392.	
XX	01-JUN-1999;	99US-0136782.	
XX	03-JUN-1999;	99US-0137222.	
XX	04-JUN-1999;	99US-0137528.	
XX	07-JUN-1999;	99US-0137502.	
XX	08-JUN-1999;	99US-0138064.	
XX	10-JUN-1999;	99US-0138540.	
XX	14-JUN-1999;	99US-0138647.	
XX	16-JUN-1999;	99US-0139119.	
XX	17-JUN-1999;	99US-0139452.	
XX	17-JUN-1999;	99US-0139453.	
XX	18-JUN-1999;	99US-0139454.	
XX	18-JUN-1999;	99US-0139455.	
XX	18-JUN-1999;	99US-0139456.	
XX	18-JUN-1999;	99US-0139457.	
XX	18-JUN-1999;	99US-0139458.	
XX	18-JUN-1999;	99US-0139459.	
XX	18-JUN-1999;	99US-0139460.	
XX	18-JUN-1999;	99US-0139461.	
XX	18-JUN-1999;	99US-0139462.	
XX	18-JUN-1999;	99US-0139463.	
XX	18-JUN-1999;	99US-0139730.	
PR	18-JUN-1999;		99US-0139763.
PR	21-JUN-1999;		99US-0139817.
PR	22-JUN-1999;		99US-0139899.
PR	23-JUN-1999;		99US-0140353.
PR	21-JUN-1999;		99US-0140354.
PR	28-JUN-1999;		99US-0140823.
PR	29-JUN-1999;		99US-0140891.
PR	30-JUN-1999;		99US-0141287.
PR	01-JUL-1999;		99US-0141842.
PR	01-JUL-1999;		99US-0142154.
PR	02-JUL-1999;		99US-0142055.
PR	06-JUL-1999;		99US-0142390.
PR	08-JUL-1999;		99US-0142803.
PR	09-JUL-1999;		99US-0142820.
PR	12-JUL-1999;		99US-0142877.
PR	13-JUL-1999;		99US-0143542.
PR	14-JUL-1999;		99US-0143624.
PR	15-JUL-1999;		99US-0144005.
PR	16-JUL-1999;		99US-0144085.
PR	16-JUL-1999;		99US-0144086.
PR	19-JUL-1999;		99US-0144325.
PR	19-JUL-1999;		99US-0144326.
PR	19-JUL-1999;		99US-0144332.
PR	19-JUL-1999;		99US-0144333.
PR	19-JUL-1999;		99US-0144334.
PR	19-JUL-1999;		99US-0144335.
PR	20-JUL-1999;		99US-0144352.
PR	20-JUL-1999;		99US-0144632.
PR	20-JUL-1999;		99US-0144632.
PR	21-JUL-1999;		99US-0144884.
PR	21-JUL-1999;		99US-0144814.
PR	21-JUL-1999;		99US-0145086.
PR	22-JUL-1999;		99US-0145088.
PR	22-JUL-1999;		99US-0145085.
PR	22-JUL-1999;		99US-0145087.
PR	22-JUL-1999;		99US-0145089.
PR	22-JUL-1999;		99US-0145192.
PR	23-JUL-1999;		99US-0145145.
PR	23-JUL-1999;		99US-0145216.
PR	23-JUL-1999;		99US-0145216.
PR	26-JUL-1999;		99US-0145276.
PR	27-JUL-1999;		99US-0145313.
PR	27-JUL-1999;		99US-0145318.
PR	27-JUL-1999;		99US-0145319.
PR	28-JUL-1999;		99US-0145951.
PR	02-AUG-1999;		99US-0146386.
PR	02-AUG-1999;		99US-0146388.
PR	02-AUG-1999;		99US-0146389.
PR	03-AUG-1999;		99US-0147038.
PR	04-AUG-1999;		99US-0147204.
PR	04-AUG-1999;		99US-0147302.
PR	05-AUG-1999;		99US-0147192.
PR	05-AUG-1999;		99US-0147260.
PR	06-AUG-1999;		99US-0147303.
PR	06-AUG-1999;		99US-0147416.
PR	08-AUG-1999;		99US-0147493.
PR	08-AUG-1999;		99US-0148375.
PR	10-AUG-1999;		99US-0148375.
PR	11-AUG-1999;		99US-0148319.
PR	12-AUG-1999;		99US-0148341.
PR	13-AUG-1999;		99US-0148565.
PR	13-AUG-1999;		99US-0148684.
PR	16-AUG-1999;		99US-0149368.
PR	17-AUG-1999;		99US-0149175.
PR	18-AUG-1999;		99US-0149426.
PR	20-AUG-1999;		99US-0149722.
PR	20-AUG-1999;		99US-0149723.
PR	20-AUG-1999;		99US-0149529.
PR	23-AUG-1999;		99US-0149502.
PR	23-AUG-1999;		99US-0149930.
PR	25-AUG-1999;		99US-0150566.
PR	26-AUG-1999;		99US-0150884.
PR	27-AUG-1999;		99US-0151065.

PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 9.7%; Score 91; DB 21; Length 650;			
Best Local Similarity 23.9%; Pred. No. 0.87;			
Matches 45; Conservative 30; Mismatches 75; Indels 38; Gaps 9;			
Qy	21	GRUDRLSKACDACSRRDPLPHFRKDSW-----PFAREASAILGLDQISHL 75	
Db	317	gvidgssxsktspaprrpppppsgesdsgplgairwatqaktsvvseuak---- 372	
Qy	76	TDELDALFDVQKARVRCGLVEDNESDSQSFPPKXKXMSNGIDNFSLDATHPG 135	
Db	373	nmkpfdcpnsitkservr--vdanq--pkpsprge--arglqrnkldettisqkn 424	
Qy	136	TP-----RSDP-----DGDALFP--CDKSKODASNDVDEELRQ--QAOQEE 174	
Db	425	ttpaanmmtrvskddafnsfvadfttdtkdgnhpkpkealeaeiqrkldeikqtksk 484	
Qy	175	ALTKAQE 182	
Db	485	aitakfe 492	
RESULT	8		

G46576	
ID	G46576 standard; Protein; 702 AA.
XX	
AC	G46576;
XX	
DF	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58611.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
FN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
XX	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132407.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132487.
PR	11-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134221.
PR	18-MAY-1999; 99US-0134370.
PR	18-MAY-1999; 99US-0134768.
PR	19-MAY-1999; 99US-0134941.
PR	20-MAY-1999; 99US-0135124.
PR	21-MAY-1999; 99US-0135353.
PR	24-MAY-1999; 99US-0135629.
PR	25-MAY-1999; 99US-0136021.
PR	27-MAY-1999; 99US-0136392.
PR	28-MAY-1999; 99US-0136782.
PR	01-JUN-1999; 99US-0137222.
PR	03-JUN-1999; 99US-0137528.
PR	04-JUN-1999; 99US-0137502.
PR	07-JUN-1999; 99US-0137724.
PR	08-JUN-1999; 99US-0138094.
PR	10-JUN-1999; 99US-0138340.
PR	14-JUN-1999; 99US-0138540.
PR	16-JUN-1999; 99US-0139149.
PR	16-JUN-1999; 99US-0139452.
PR	17-JUN-1999; 99US-0139453.
PR	18-JUN-1999; 99US-0139454.
PR	18-JUN-1999; 99US-0139455.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 21-JUN-1999; 99US-0140033.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 23-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143244.  
 PR 13-JUL-1999; 99US-0143244.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144332.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151436.  
 PR 01-SEP-1999; 99US-0151436.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 06-OCT-1999; 99US-0157423.  
 PR 06-OCT-1999; 99US-0157423.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159884.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160810.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161520.  
 PR 28-OCT-1999; 99US-0161592.  
 PR 28-OCT-1999; 99US-0161593.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 9.7%; Score 91; DB 21; Length 702;  
 Best Local Similarity 23.9%; Pred No. 0.97; 75; Indels 38; Gaps 9;  
 Matches 45; Conservative 30; Mismatches 30; Mismatches 38; Gaps 9;  
 QY 21 GRDRTLSKAGSDAGDSRDLPHFRGKSDW-----PYARREASAILGLDQKISHL 75  
 Db 317 gvhdskaatkpsapaprrpppppsgesdgpglgaifwathqaktsvaednk----- 372  
 QY 76 TDDELDFDQVAFARVGLVEDNEDSDSQSPFRDYKYNMGITDNFSLDTAITMFG 135  
 Db 373 nmpktdpnstsktsrver---vdsdq---pktpspvrge--arglqrnkldettlsqkn 424  
 QY 136 TP-----RSDG-----DGDALFF--GDKSKQDSNVDVEELRQ--OAOEMEE 174  
 Db 425 ttpaannmtrvskdafsnafvadftdtktdngnkpdkaealeaigrlkdelqtksek 484  
 QY 175 ALTKAQE 182

Db 485 aeltakfe 492  
1 111 1  
RESULT 9  
G15482  
ID G15482 standard; Protein; 725 AA.  
AC G15482;  
XX  
XX  
XX 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 15751.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000BP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147036.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147322.  
PR 05-AUG-1999; 99US-0147306.  
PR 05-AUG-1999; 99US-0147307.  
PR 06-AUG-1999; 99US-0147309.  
PR 06-AUG-1999; 99US-0147316.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.

PR	17-AUG-1999:	99US-0149175.	OY	136 TP-----RSD-----GDALFF--GDKSKODASNVDEELROO--OQAMEE 174
PR	18-AUG-1999:	99US-0149426.	DB	500 ttpaaanmrsvskddafnsfvadfttkfdngnkgpkgaaleaeigrkldeikqtksk 559
PR	20-AUG-1999:	99US-0149722.	OY	175 ALKTAOE 182
PR	20-AUG-1999:	99US-0149723.	DB	560 aaitakfe 567
PR	20-AUG-1999:	99US-0149929.	DB	
PR	23-AUG-1999:	99US-0149902.		
PR	23-AUG-1999:	99US-0149930.		
PR	23-AUG-1999:	99US-0150566.		
PR	26-AUG-1999:	99US-0150884.		
PR	26-AUG-1999:	99US-0151062.		
PR	27-AUG-1999:	99US-0151063.		
PR	27-AUG-1999:	99US-0151060.		
PR	30-AUG-1999:	99US-0151100.		
PR	30-AUG-1999:	99US-0151103.		
PR	31-AUG-1999:	99US-0151438.		
PR	01-SEP-1999:	99US-0151930.		
PR	07-SEP-1999:	99US-0152363.		
PR	10-SEP-1999:	99US-0153070.		
PR	13-SEP-1999:	99US-0153758.		
PR	15-SEP-1999:	99US-0154018.		
PR	16-SEP-1999:	99US-0154039.		
PR	20-SEP-1999:	99US-0154779.		
PR	22-SEP-1999:	99US-0155139.		
PR	23-SEP-1999:	99US-0155486.		
PR	24-SEP-1999:	99US-0155659.		
PR	28-SEP-1999:	99US-0156458.		
PR	29-SEP-1999:	99US-0156759.		
PR	04-OCT-1999:	99US-0157159.		
PR	05-OCT-1999:	99US-0157743.		
PR	06-OCT-1999:	99US-0157765.		
PR	07-OCT-1999:	99US-0158029.		
PR	08-OCT-1999:	99US-0158232.		
PR	12-OCT-1999:	99US-0158369.		
PR	13-OCT-1999:	99US-0159293.		
PR	13-OCT-1999:	99US-0159294.		
PR	13-OCT-1999:	99US-0159295.		
PR	14-OCT-1999:	99US-0159329.		
PR	14-OCT-1999:	99US-0159330.		
PR	14-OCT-1999:	99US-0159331.		
PR	14-OCT-1999:	99US-0159637.		
PR	14-OCT-1999:	99US-0159638.		
PR	18-OCT-1999:	99US-0159584.		
PR	21-OCT-1999:	99US-0160741.		
PR	21-OCT-1999:	99US-0160756.		
PR	21-OCT-1999:	99US-0160770.		
PR	21-OCT-1999:	99US-0160814.		
PR	21-OCT-1999:	99US-0160815.		
PR	22-OCT-1999:	99US-0160980.		
PR	22-OCT-1999:	99US-0160981.		
PR	22-OCT-1999:	99US-0160989.		
PR	25-OCT-1999:	99US-0161404.		
PR	25-OCT-1999:	99US-0161405.		
PR	25-OCT-1999:	99US-0161406.		
PR	26-OCT-1999:	99US-0161359.		
PR	26-OCT-1999:	99US-0161360.		
PR	26-OCT-1999:	99US-0161361.		
PR	28-OCT-1999:	99US-0161920.		
PR	28-OCT-1999:	99US-0161924.		
PR	28-OCT-1999:	99US-0161925.		
PR	28-OCT-1999:	99US-0162172.		
PR	29-OCT-1999:	99US-0162142.		
Query Match 9.7%: Score 91; DB 21; Length 725;				
Best Local Similarity 23.9%: Pred No. 1:				
Matches 45; Conservative 30; Mismatches 75; Indels 38; Gaps 9;				
OY	21 GRHRTLSKAGSDGDSOSPLPIFRGKDSW-----FVAREAAASAILGLDQKISHL 75			
DB	392 gvhgskskskpseprtspppppsgesdggplgafwaqhktsvvsednk--- 447			
OY	75 TDDELALDFQKARAVRGLVEVNEEDSSQSFVRDKYMSNGTIDNFSLDIAITMPG 135			
DB	448 nmpkfdepnantakservr---vdahq---pkpsspyrge--ardgrnkdtlettsckn 499			
RESULT 10 G35664 standard; Protein; 326 AA. ID G35664; AC G35664; XX DT 18-OCT-2000 (first entry) DE Arabidopsis thaliana protein fragment SEQ ID NO: 43602. XX KW Protein identification; signal transduction pathway; metabolic pathway; KW hybridisation assay; genetic mapping; gene expression control; promoter; KW termination sequence. XX OS Arabidopsis thaliana. XX XN EPI033405-A2. PN PR 06-SEP-2000. XX PF 25-FEB-2000; 2000EP-0301439. XX PR 25-FEB-1999; 99US-0121825. PR 05-MAR-1999; 99US-0123180. PR 09-MAR-1999; 99US-0123548. PR 23-MAR-1999; 99US-0125788. PR 25-MAR-1999; 99US-0126264. PR 29-MAR-1999; 99US-0126785. PR 01-APR-1999; 99US-0127462. PR 06-APR-1999; 99US-0128234. PR 08-APR-1999; 99US-0128714. PR 16-APR-1999; 99US-0129845. PR 19-APR-1999; 99US-0130077. PR 21-APR-1999; 99US-0130449. PR 21-APR-1999; 99US-0130540. PR 23-APR-1999; 99US-0130990. PR 28-APR-1999; 99US-0131449. PR 30-APR-1999; 99US-0132048. PR 30-APR-1999; 99US-0132407. PR 04-MAY-1999; 99US-0132484. PR 05-MAY-1999; 99US-0132485. PR 06-MAY-1999; 99US-0132486. PR 07-MAY-1999; 99US-0132487. PR 07-MAY-1999; 99US-0132863. PR 11-MAY-1999; 99US-0134256. PR 11-MAY-1999; 99US-0134218. PR 14-MAY-1999; 99US-0134219. PR 14-MAY-1999; 99US-0134221. PR 14-MAY-1999; 99US-0134370. PR 16-MAY-1999; 99US-0134766. PR 18-MAY-1999; 99US-0135121. PR 20-MAY-1999; 99US-0135121. PR 21-MAY-1999; 99US-0135353. PR 24-MAY-1999; 99US-0135629. PR 25-MAY-1999; 99US-0136021. PR 27-MAY-1999; 99US-0136392. PR 28-MAY-1999; 99US-0136782. PR 01-JUN-1999; 99US-0137222. PR 03-JUN-1999; 99US-0137528. PR 04-JUN-1999; 99US-0137502. PR 07-JUN-1999; 99US-0137724. PR 08-JUN-1999; 99US-0138094. PR 10-JUN-1999; 99US-0138540. PR 10-JUN-1999; 99US-0138847.				

[illegible]

```
QY 12 NSQLGSPAPGRHDR-TLSKAGSDADGDSRS-----DSPLPHFRGKDSDF 55
```

QY 15 LGSPAPGRHD--RTLSKAGSDADGDSRSDSPLPHFRGKDSDFWYARRZAASAILGLDQK 71

309	lspspsphdfikawryrgpsteddd	-----dsdwsaeeg-----ka 34
72	ISHLTDELDALFDVOKARAVRELVDNEDSDSSFFPVRYKMSYCNIDNFS	-----126
349	lspspsphdfi-----kawyrg-edtd-dqsdwgsaek-----dglagtfaiphts 397	
127	--LUTATWPGTSPSSDDGDALFFGDKKODASNDVVEELRQQQQA 171	
398	afktwccpstedtddecevv-----pdseaadpkpsheaq 438	
13	RESULT	
R41044	ID	R41044 standard; protein; 1028 AA.
R41044	XX	
R41044	AC	
22-MAR-1994	DT	(first entry)
XX	DT	
DE	DE	Plasmodium vivax Duffy receptor residues 23-1051.
KW	KW	Merozoite; malaria; HIV; env; CD4; truncated Duffy receptor;
KW	KW	human immunodeficiency virus; envelope glycoprotein; hybrid protein;
KW	KW	red blood cell; erythrocyte; AIDS; cross-linking agent;
KW	KW	joining region.
XX	XX	
OS	OS	Plasmodium vivax.
XX	XX	
Key	Key	Location/Qualifiers
FF	FF	1028
FF	FF	Modified-site
FF	FF	/note- "opt. terminates with a Cys residue for
FF	FF	disulphide bond formation with an added Cys
FF	FF	at the C-terminal end of a CD4 fragment;
FF	FF	alternatively, Cys residues are
FF	FF	incorporated near the C-terminal end of
FF	FF	P.vivax Duffy 23-1051 and of CD4 1-371 to
FF	FF	allow disulphide cross-linking or the two
FF	FF	parts of the fusion protein could be
FF	FF	connected via any linking molecule or agent."
PN	PN	W09318160-A.
XX	XX	
PD	PD	16-SEP-1993.
XX	XX	
PF	PF	10-MAR-1993; 93WO-GB00505.
XX	XX	
PR	PR	11-MAR-1992; 92GB-0005276.
PR	PR	08-JUL-1992; 92GB-0014461.
PR	PR	24-JUL-1992; 92GB-0015829.
PR	PR	14-JUL-1992; 92GB-0015962.
PR	PR	03-MAR-1993; 93GB-0004311.
XX	XX	(PREN/) PRENDERGAST K F.
XX	XX	
PT	PT	Prendergast KF;
XX	XX	
WPI	WPI	1993-30347/38.
DR	DR	
PT	PT	Anti-viral fusion peptide(s) - comprise viral-binding component
PT	PT	and malaria merozoite red cell binding component, for treating
PT	PT	e.g. HIV, and hepatitis
PS	PS	
PS	PS	Claim 11; Page 49-51; 69pp; English.
XX	XX	
CC	CC	A hybrid protein in which the P.vivax Duffy receptor is joined to
CC	CC	the CD4 receptor molecule, both molecules being truncated at their
CC	CC	transmembrane domains, is specifically claimed. The fusing or
CC	CC	joining of the two segments takes place by joining amino acid
CC	CC	segments at unspecified sites by disulphide bonds or by cross-linking
CC	CC	agents of any type. The fusion protein can bind free HIV in the blood
CC	CC	to red blood cells and consequently reduce viral titre, prevent
CC	CC	transmission of the virus and improve safety of blood transfusions.
CC	CC	The specification also includes the sequence of amino acids 1-371



PT New malaria vaccines - contains cysteine-rich DBL family protein  
PT binding domains homologous domains of the Duffy and sialic acid  
PT binding proteins

Search completed: April 25, 2001, 10:12:50  
Job time: 346 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 25, 2001, 10:13:24 ; Search time 31.19 Seconds  
(without alignments)  
112.715 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_602\_784

Perfect score: 938

Sequence: 1 QROSLRHVSINQLGSPAP.....ELRQQQAWEALKTAQKEF 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	89.5	9.5	590	3 US-08-893-852A-4
2	89.5	9.5	590	4 US-08-821-818-2
3	89.5	9.5	1115	2 US-08-568-459A-2
4	89.5	9.5	1115	2 US-08-487-826B-2
5	89.5	9.5	1115	6 5198347-6
6	87.5	9.3	933	3 US-08-293-728-2
7	87.5	9.3	933	4 US-08-421-868-2
8	85.5	9.1	1820	3 US-07-998-289B-8
9	85.5	9.1	2100	2 US-08-608-793-23
10	85.5	9.1	2100	3 US-08-772-512A-19
11	80	8.5	1087	1 US-08-264-002-5
12	78	8.3	3052	2 US-08-557-122A-26
13	76.5	8.2	200	2 US-08-531-525-33
14	76.5	8.2	200	2 US-08-718-270A-33
15	76.5	8.2	1202	4 US-08-576-235-14
16	73.5	8.0	780	2 US-08-488-801-1
17	73.5	8.0	780	2 US-08-973-831-2
18	73.5	8.0	780	2 PCT-US96-09510A-2
19	75.5	8.0	454	3 US-08-348-518C-4
20	75	8.0	454	3 US-08-476-509B-4
21	75	8.0	454	3 US-08-808-793-3
22	74	7.9	2105	2 US-08-772-512A-3
23	74	7.9	2105	3 US-08-531-525-32
24	73.5	7.8	204	2 US-08-718-270A-32
25	73.5	7.8	889	5 PCT-US93-11725-2
26	73.5	7.8	889	5 PCT-US93-11725-2
27	73.5	7.8	4550	2 US-08-804-227C-8

28	73.5	7.8	4550	2 US-08-804-198-2
29	72.5	7.7	326	2 US-08-757-653-172
30	72.5	7.7	326	2 US-08-823-516-75
31	72.5	7.7	326	2 US-08-823-516-135
32	72.5	7.7	326	3 US-08-759-038-111
33	72.5	7.7	326	3 US-08-758-314-111
34	72.5	7.7	380	2 US-09-026-587-4
35	72.5	7.7	380	2 US-09-227-420-4
36	72	7.7	764	1 US-08-375-300-4
37	72	7.7	764	1 US-09-177-431-4
38	72	7.7	764	5 US-08-375-300-4
39	72	7.7	1089	1 US-08-973-831-2
40	72	7.7	1089	3 US-09-177-431-2
41	72	7.7	1089	5 PCT-US95-16430-2
42	72	7.7	1349	4 US-08-938-291A-6
43	71.5	7.6	249	1 US-08-466-603-2
44	71.5	7.6	249	1 US-08-314-503A-2
45	71.5	7.6	249	1 US-08-468-066-2

#### ALIGNMENTS

RESULT 1  
US-08-893-852A-4  
; Sequence 4, Application US/08893852A  
; Patent No. 5080558  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; INVENTOR: Shih, Peck  
; APPLICANT: Shih, Peck  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: NEW HUMAN GROWTH REGULATOR PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; CURRENT APPLICATION NUMBER: FastUSQ for Windows Version 2.0  
; APPLICATION NUMBER: US/08/893,852A  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0341 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 590 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 452490  
US-08-893-852A-4

Query Match 9.5%; Score 89.5; DB 3; Length 590;  
Best Local Similarity 25.7%; Pred. No. 0.15;

Matches 43; Conservative 17; Mismatches 60; Indels 47; Gaps 9;

```
QY 15 LGSPAPGRHD---RTLKAGSDADGDSKSDPLPHFRGKDSDFYARREASAILGLDQK 71
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 309 LSSPTSPHDFLKAWVYRPGEDTDDO-----DSDWGSAAEEG-----KA 348
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 72 TSHLTDELDAFDVQKARAVRGLVEDNEDSQSPFVRDKYMSNGTIDNFS----- 126
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 349 LSSPTSPHDFL-----KAWVYRPG--EDTED-DQSDWGSAAEK---DGLAOTFAIPTS 397
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 127 --LDTAITMPTGRSDDGDAFFGDKKSKQDASNVDFELRQOQAO 171
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 398 AFLKTWCCPGEDTDDCEVVV-----PEDEAANDPKSPSHEAQ 438
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

## RESULT 2

```
US-08-821-818-2
; Sequence 2, Application US/08821818
; Patent No. 6146877
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; TITLE OF INVENTION: IDENTIFICATION OF THE PROGRESSION
; TITLE OF INVENTION: ELEVATED GENE-3 AND USES THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,818
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Chag, Albert Wal-Kit
; REGISTRATION NUMBER: 36,479
; REFERENCE/DOCKET NUMBER: 51523
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 478-0400
; TELEFAX: (212) 351-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGIN:
; ORGANISM: US-08-821-818-2
```

Query Match 9.5%; Score 89.5; DB 4; Length 590;

```
Best Local Similarity 25.7%; Pred. No. 0.15;
Matches 43; Conservative 17; Mismatches 60; Indels 47; Gaps 9;

QY 15 LGSPAPGRHD---RTLKAGSDADGDSKSDPLPHFRGKDSDFYARREASAILGLDQK 71
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 309 LSSPTSPHDFLKAWVYRPGEDTDDO-----DSDWGSAAEEG-----KA 348
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 72 TSHLTDELDAFDVQKARAVRGLVEDNEDSQSPFVRDKYMSNGTIDNFS----- 126
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 349 LSSPTSPHDFL-----KAWVYRPG--EDTED-DQSDWGSAAEK---DGLAOTFAIPTS 397
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 127 --LDTAITMPTGRSDDGDAFFGDKKSKQDASNVDFELRQOQAO 171
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 398 AFLKTWCCPGEDTDDCEVVV-----PEDEAANDPKSPSHEAQ 438
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

## RESULT 3

```
US-08-568-459A-2
; Sequence 2, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhuan
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1115 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGIN:
; ORGANISM: Plasmodium vivax
; US-08-568-459A-2
```

Query Match 9.5%; Score 89.5; DB 2; Length 1115;

```
Best Local Similarity 24.0%; Pred. No. 0.39;
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;

QY 6 LRHSVNTSQGLSPAPGRHDTLSKAGSDADGDSRSDSPLPHFRGKDSDFYARREASAI 65
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 716 LRHSKNDSDSGFA-----ESMANPDSNKGEGT-----KQDNDMAKATKDSNS 762
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 66 LGIDQKISHITDDELDAFDVQKARAVRGLVEDNEDSDSFFVRDKYMSN---GTI 122
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 763 DG-----TSATGGTTDAV-----DREINRGVPED-----RKIVGSKDGGGE 800
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 123 DNFSLDTAITMPTGRSDDGDAFFGDKKSKQDASNVDFE---LRQOQAMEALKA 179
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 801 DNANKADKATVVGEDIRIENSAG-----GSTNDRSKNUTERNKASTPQSKSEDATLS 854
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 180 KQE 182
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 855 KTE 857
      |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

## RESULT 4

```

; APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,
; DAVID C.; FANG, XIANGDOU
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
; PLASMODIUM KNOWLESII DUFFY RECEPTOR
; NUMBER OF SEQUENCES: 27
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554, 837
; FILING DATE: 20-JUL-1990
; SEQ ID NO:6:
; LENGTH: 1115
519347-6

Query Match          9.5%; Score 89.5; DB 6; Length 1115;
Best Local Similarity 24.0%; Pred. No. 0.39;
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;

QY 6 LRHSVTSQSGAPGCRHDIRTLKSAGSDADGDSRSOSPLPHRGKDSDFWYARREASAT 65
DB 716 LRHSKDNSSDGA-----ESMANPDSKSGETG---KGDNDMAKATKDSNS 762

QY 66 LGLDQKISHLTDDLDALFDVOKARVRGLVEDNDSQSSPPVROKYMKN---GTI 122
DB 763 DG-----TSSAGTGTDAV-----DREINKGVPED-----RKYVSGKDGG 800

QY 123 DNEISLDTATWPTGTRSDSDGDLFGDKSKQASNDVVE-----LQQQQAWEALKTA 179
DB 801 DNSAKNDATVVGEDIRENSAG-----GSTNDRSKNPTKMGASTPQSQEDATLS 854

QY 180 KDE 182
DB 855 KTE 857

RESULT 6
US-08-293-728-2
; Sequence 2, Application US/08293728D
; Patent No. 6008341
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; INVENTOR: Foster, Timothy J.; Baren, L.
; TITLE OF INVENTION: Plasmodium Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/08/293,728D
; CURRENT FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-293-728-2

Query Match          9.3%; Score 87.5; DB 3; Length 933;
Best Local Similarity 25.6%; Pred. No. 0.49;
Matches 42; Conservative 18; Mismatches 61; Indels 43; Gaps 5;

QY 9 SVTNSQSGAPGCRHDIRTLKSAGSDADGDSRSOSPLPHRGKDSDFWYARREASATLGI: 68
DB 795 SDSDSDSDSDSSESD-----SDSESDSDSDSDSDSDSDSDSDS-----SASDSDSGS 845

QY 69 DQKISHLTDDLDALFDVOKARVRGLVEDNDSQSSSF-----PVDRKYMKNGTIND 124
DB 846 DSDSDSDSDSDSDSDS-----NSDSGSGSNMNVVPPNPGKNGTNASKN 887

QY 125 FSLDTATWPTGTRSDSD-----GDLPTGDKSKQD 156
DB 888 EAKDSKEPTPTGSGEDANTSLWGLLASIGSLLFRKKNKD 931

RESULT 7
US-09-409-421-868-2

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-289B-8

Query Match          9.38; Score 87.5; DB 4; Length 933;
Best Local Similarity 25.68; Pred. No. 0.49;
Matches 42; Conservative 18; Mismatches 61; Indels 43; Gaps 5;

QY 9 SVTNSQLGSPAPGRDRTLSKAGSDADGSDSPPLPHFRGKSDWFIYARREASAILGL 68
Db 795 SDSQSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSGS 845
QY 69 DKSHITDDELDALEDDYQKARAVRGVLVDNEDSDSQSF---PVRYKYSNGTIDN 124
Db 846 SDSQSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSGS 887
QY 125 FSLDTATMPGTPRSDDD-----GDALFFGDKSKOD 156
Db 888 EKASKEPLPTDGSDEDAWNTSLWGLLASIGLLFLFRKKENKD 931

RESULT 8
US-07-998-289B-8
; Sequence 8, Application US/07998289B
; Patent No. 6028746
; GENERAL INFORMATION:
; APPLICANT: Black, Bruce C
; APPLICANT: Taylor, Haric
; APPLICANT: Hays, David G
; TITLE OF INVENTION: Method for Monitoring Pesticide
; TITLE OF INVENTION: Resistance
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07998,289B
; FILING DATE: 30-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Robinson, Joseph R
; REGISTRATION NUMBER: 33,448
; REFERENCE/DOCKET NUMBER: 0646/OA939
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-289B-8

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-289B-8

Query Match          9.1%; Score 85.5; DB 3; Length 1820;
Best Local Similarity 21.9%; Pred. No. 2.1;
Matches 44; Conservative 36; Mismatches 70; Indels 51; Gaps 9;

QY 4 SLLRHSVTNSQLGSPAPGRDRTLSKAGSDADGSDSPPLPHFRGKSDWFIYARREAS 63
Db 1016 ALLSNFGSSLSAPT-----ADNTNKIAFAFNIGRFSM--VKRNIA 1059
QY 64 AILGLDQKI-----SHLTDELDALPDYQKARAVRGVLVDNE----- 102
Db 1060 CPKLRNKLNTQISDPSHGDNLELGHDEILDGLIKKIKGKEOTOLEVALDGMFTI 1119
QY 103 DSDSQSPFVRKYMSNGTIDNFSLDTA-TMPGTPRSDDDGALFFG-----DKSKOD 156
Db 1120 HGDMMNNKPKSKYLNATDD-----DTASINSYGHKNRPFKDSHKSAGTSMGEEKR 1175
QY 157 ASNVDV---BELRQOQAQME 174
Db 1176 ASKEDLGLDEL-DEGECEE 1195

RESULT 9
US-08-808-793-23
; Sequence 23, Application US/08808793
; Patent No. 5858713
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,793
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Braman, Susan J
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-289B-8

```

```

; LENGTH: 2100 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-772-512A-19

Query Match          9.1%; Score 85.5; DB 2; Length 2100;
Best Local Similarity 21.9%; Pred. No. 2.6;
Matches 44; Conservative 36; Mismatches 70; Indels 51; Gaps 9;

OY 4 SILRHSTVNSQLSGSPAPGRDRTLTKAGSDADGSDSDSLPHFRKQSDHFWYARREAS 63
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1029 ALLSNFGSSLSAPT-----AONDNTKIAEAFNRIGREKSW--VKRNIA 1072

OY 64 ATILGDKIT-----SHLTDELDALFDVOKARAVRGLVEDNE----- 102
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1073 CFKLIRNKLTNQISDQSEHGDNLELGHDEILADGLIKGKEQTOLVAIGDMFTI 1132

OY 103 DSDSQSFPVDRKYMSNGTIDNFSLDTA-ITMPTGPRSDDGDALEFFG-----DKSKQD 156
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1133 HGMKNKPKSKYLNATDD---DTASINSYGSHKRNPFKDESHKGSATMETGECKRD 1188

OY 157 ASKVDV---EELRQQAQNEE 174
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1189 ASKEDGLDEEL-DEEGECE 1208

RESULT 10
US-08-772-512A-19
; Sequence 19, Application US/08772512A
; Patent No. 6022705
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON HARGRAVE, DEVANS & DOYLE LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772-512A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,618
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brame, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/601(CRFD-1657)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2100 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
; US-08-772-512A-19

Query Match          9.1%; Score 85.5; DB 3; Length 2100;
Best Local Similarity 21.9%; Pred. No. 2.6;
Matches 44; Conservative 36; Mismatches 70; Indels 51; Gaps 9;

OY 4 SILRHSTVNSQLSGSPAPGRDRTLTKAGSDADGSDSDSLPHFRKQSDHFWYARREAS 63
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1029 ALLSNFGSSLSAPT-----AONDNTKIAEAFNRIGREKSW--VKRNIA 1072

OY 64 ATILGDKIT-----SHLTDELDALFDVOKARAVRGLVEDNE----- 102
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1073 CFKLIRNKLTNQISDQSEHGDNLELGHDEILADGLIKGKEQTOLVAIGDMFTI 1132

OY 103 DSDSQSFPVDRKYMSNGTIDNFSLDTA-ITMPTGPRSDDGDALEFFG-----DKSKQD 156
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1133 HGMKNKPKSKYLNATDD---DTASINSYGSHKRNPFKDESHKGSATMETGECKRD 1188

OY 157 ASKVDV---EELRQQAQNEE 174
   || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1189 ASKEDGLDEEL-DEEGECE 1208

RESULT 11
US-08-264-002-5
; Sequence 5, Application US/08264002
; Patent No. 5558019
; GENERAL INFORMATION:
; APPLICANT: GUI, XIANG-FANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUNARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1087 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: CEHK
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1087
; US-08-264-002-5

Query Match          8.5%; Score 80; DB 1; Length 1087;

```

RESULT 14  
US-08-718-270A-33  
; sequence 33, Application US/08718270A  
; Patent No. 5910478  
; GENERAL INFORMATION:  
; APPLICANT: Hlavka, Joseph J.



Query Match	8.2%	Score 76.5	DB 4	Length 1503
Best Local Similarity	24.7%	Pred No. 16		
Matches	48	Conservative	19	Mismatches 64
			Indels	Gaps
QY	10	VTNISGLSPGPGHRTLTSLKAGSDGDSRSDPL-----PHFGRKDSDFVARR	59	
DB	509	VFESLSLDRGEGKQ-----DSGQDVLPGVGPVDFAHNLVSVDGYIQL	556	
QY	60	EASAIAIGLDGDKISHLTDD-----ELDAFDVOKARVGLVEDNEEDSD-S	106	
DB	557	EXSGNLELDYYPALLTITDMNMPRTGPELSQIAL-----RSVE--LBERSDTEDFP	607	
QY	107	QSSFPVDKYSXNGTILNLSLDATIM--PQTRSDSDGDLFTGDKSKSQASN----	161	
DB	608	QSSTDPKDSI.PG-----DLHVTSGPSPFNK-----IFNDVKSEDLFSHQIFD	653	
QY	162	VEELRQQQAQMEEA	175	
DB	654	LMELNGVQADEKPA	667	

RESULT 15  
US-08-976-255-14  
; Sequence 14, Application US/08976255  
; Patent No. 6136581  
; GENERAL INFORMATION:  
; APPLICANT: Jono, Keith E.  
; APPLICANT: Plowman, Gregory

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:14:12 ; Search time 46.78 Seconds  
(without alignments)  
266.839 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_602\_784

Perfect score: 938

Sequence: 1 QESLRLHSVTNSQLSPAP.....ELRQQAQMEALTKAQEF 183

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database: PIR.67.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	10.7	2649	2 T51023	hypothetical prote
2	98.5	10.5	1883	2 T13944	chromodomain-helic
3	91.5	10.0	599	2 T39990	hypothetical prote
4	91.5	10.0	1254	2 T41262	mutS family DNA m
5	91.5	9.8	3839	2 T49799	related to TOM1 pr
6	90.5	9.6	1707	2 T18951	hypothetical prote
7	90	9.6	611	2 T06458	nucleolin homolog
8	89.5	9.5	590	2 A56535	gad434 protein - 1
9	89.5	9.5	1070	2 T08484	Duffy receptor - P
10	89	9.5	476	2 T36395	probable transcrip
11	88.5	9.4	1353	2 T00347	hypothetical prote
12	88	9.4	384	2 S51796	vasodilator-stimul
13	87.5	9.3	233	2 A36425	clathrin light cha
14	87.5	9.3	1302	2 S41539	surface-located me
15	87	9.3	1302	2 J60009	hypothetical prote
16	86.5	9.2	1937	2 T38077	hypothetical coile
17	86	9.2	1937	2 T38077	hypothetical coile
18	85.5	9.1	438	2 B39487	transcription init
19	85.5	9.1	629	2 B64075	hypothetical prote
20	85.5	9.1	805	2 T06557	sodium channel pro
21	85.5	9.1	1820	2 A33299	hypothetical prote
22	85.5	9.1	2946	2 T00867	hypothetical prote
23	85	9.1	747	2 T33488	85K c-Cbl-interact
24	84	9.0	665	3 JC7191	forA protein precu
25	83.5	8.9	415	2 S35760	hypothetical prote
26	83.5	8.9	438	2 S13637	chromosomal protei
27	83.5	8.9	1203	2 B55094	hypothetical prote
28	83.5	8.9	1304	2 T19397	hypothetical prote
29	83	8.8	732	2 T32757	hypothetical prote

apaB protein - Eme  
hypothetical prote  
t-complex protein  
hypothetical prote  
hypothetical prote  
hypothetical prote  
enhancer-trap-locu  
hypothetical prote  
hypothetical prote  
DNA-binding domain  
ferric-pseudobacti  
cell surface glyco  
acidic nuclear pro  
nucleolin homolog  
excision repair pr

## ALIGNMENTS

RESULT 1  
T51023  
hypothetical protein B7F21.40 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T51023 V: Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu  
submitted to the Protein Sequence Database, July 2000  
A:Reference number: 225286  
A:Accession: T51023  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-2649 <SCH>  
A:Cross-references: EMBL:AL389901; GSPDB:GN00116; NCSP:B7F21.40  
A:Experimental source: BAC clone B7F21; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B7F21.40  
A:Map position: 6  
A:Introns: 1619/3; 2584/1

Query Match  
Res Local Similarity 10.7% Score 100; DB 2; Length 2649;  
Matches 54; Conservative 29; Mismatches 80; Indels 52; Gaps 10;  
QY 10 VTNQLGSPAG-----RHDTLTKAGSD-----ADGDSRS-----DSP 43  
DB 559 VSQSQRGRFGHPHSPSHDHGCTERRPRSPDKASESHLVAAQGSRSLLTGANDTP 618  
QY 44 LPHFGRKDSMFARREAAALGLDKQI-----SHLTDELDALFDQVKARVRGLV 98  
DB 619 IRPFRGTQSARASVDRDTRASLDNRMDRTDSFGGAHRSFEIDA---PRDRERPAIMAPM 675  
QY 99 EDNEDSDS-----QSSFPVPRDK--YNSNETIDNFSLDTAIT--MPGTPRSDODGDALFFGD- 150  
DB 676 RQEDADKGRIPSRVPRKRSVLALPSTRQLPQTQTSULPT-ESDDEMDMDYFDA 734  
QY 151 -----KSKQDSQSWVDVELRQQQKMEAL 176  
DB 735 ETRKQETELKLMKDSACVPMQIVQVACTVHDNM 769

RESULT 2  
T13844  
chromodomain-helicase-DNA-binding protein - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
C:Accession: T13844  
R:Stokes, D.G.; Tartof, K.D.; Perry, R.P.  
Proc. Natl. Acad. Sci. U.S.A. 93, 7137-7142, 1996  
A:Title: CHD1 is concentrated in interbands and puffed regions of Drosophila polytene  
A:Reference number: 217823; MUID: 96293489

A:Accession: T13944  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1883 <STO>  
 A:Cross-references: EMBL:L77907; NID:g1448982; PID:g1448983; PIDN:AAC37264.1  
 C:Genetics:  
 A:Gene: CHD-1  
 A:Cross-references: FlyBase:FBgn0016132  
 C:Superfamily: CHD-1 protein; chromobox homology  
 C:Keywords: DNA binding

Query Match 10.5%; Score 98.5; DB 2; Length 1883;  
 Best Local Similarity 23.1%; Pred. No. 5.7;  
 Matches 42; Conservative 24; Mismatches 73; Indels 43; Gaps 6;

Qy 8 HSYNSQLGSPAPGRHRTLSKAGSDAGDSRS-----DSPLPFRGKDSDFYARRE 60

Db 29 HSGSGSGSGS-----SGSDSDSDSSGSGRSPDPKSLSVAGFPPTAA 76

Qy 61 AASAILGLDQIKISHLTDELDALFDVQKARAVRGLVNEEDSDSQSFVRDKYMSNG 120

Db 77 AAQA-----DKTNGFTDQDSSDG-----SSGSDSDSDAEGFSDQR---NQ 117

Qy 121 TIDNFSLDTAITMPTGTPRSDDGDLFFGDKSKODASNDVVEELRQQAQMEALKTAK 180

Db 118 SINANTSSSLPKPEQNEEDNET-----EAGQQQAPASDADESSDSANYSPTSSSS 172

Qy 181 QE 182

Db 173 SE 174

Query Match 10.0%; Score 93.5; DB 2; Length 1254;  
 Best Local Similarity 20.5%; Pred. No. 8.9;  
 Matches 38; Conservative 33; Mismatches 71; Indels 43; Gaps 5;

Qy 12 NSQLGSPAPGRHRTLSKAGSDAGDSRSPLPFRGKDSDFYF---RREAASAILG 67

Db 121 HSLGSESPGKLLRTSVKQPDSEEDSPTKKSKSLDTSIFQADQFRHPVSSKLEN 180

Qy 68 LD-----OKTSHLTDELDALFDVQKARAVRGLVNEEDSD----- 105

Db 181 SELSEVDKPPFTASRRSRKRPVSAESDE-DEDFDAPTKGSRHKRIVSDSDSDVPEPHI 239

Qy 106 ----SSQSFVRDKYMSNGTIDNFSLDTAITMPTGTPRSDDGDLFFGDKSKODASNDV 161

Db 240 SEASSEASLPIDEVSMDEDDVGYSDHSVSAAPIPK-----KSRKSSNSL 287

Qy 162 VEELR 166

Db 288 YESYR 292

Query Match 10.0%; Score 93.5; DB 2; Length 599;  
 Best Local Similarity 19.8%; Pred. No. 3.7;  
 Matches 52; Conservative 29; Mismatches 72; Indels 109; Gaps 9;

Qy 8 HSYNSQLGSPAPGRHRTLSKAGSDAGDSRSPDL-----PFRGKDSDFYARREA 61

Db 217 HQSFNSLITQPTTYNFRNFESINDASSDSDAPLRLTSSPSRLRMKNDKRYLVEHS 276

Qy 62 ASAIL-----GLDQK----- 71

Db 277 PAALIKETSIDGDKSLASSTREVSVPNEEDSVNDDSDSDVSEKETEAKHEIRAPA 336

Qy 72 -----ISH-----LTDELDAL-----FDVQKARAVRGL 97

Db 337 IIVRETSSHPSTAVPSENDTETESNDTLSESTWTSISSPSPNSDSDTLTKVDSPNKSL 396

Qy 98 VEEN-----EDSDSQSFVRDKYMSNGTIDNFSLDTAITMPTGTPRSDDGDLFFGD 150

Db 397 VDNVNSAKHKEGSKGKPPPPSQTTLV-----TSTISAAGNPSDEIGSE---ND 445

Qy 151 KKSQODA-SNVDVEELRQQAQ 171  
 Db 446 SDSDSDDSDSVPLSLQKKSQ 467

# RESULT 4

T41262  
 multis family DNA mismatch repair protein - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T41262  
 R:Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z21981  
 A:Accession: T41262  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1254 <SEE>  
 A:Cross-references: EMBL:AL031545; PIDN:CAA20855.1; GSPDB:GN00068; SPDB:SPCC285.16c  
 C:Genetics:  
 A:Gene: SPDB:SPCC285.16c  
 A:Map position: 3  
 A:Introns: 198/3

Query Match 10.0%; Score 93.5; DB 2; Length 1254;  
 Best Local Similarity 20.5%; Pred. No. 8.9;  
 Matches 38; Conservative 33; Mismatches 71; Indels 43; Gaps 5;

Qy 12 NSQLGSPAPGRHRTLSKAGSDAGDSRSPLPFRGKDSDFYF---RREAASAILG 67

Db 121 HSLGSESPGKLLRTSVKQPDSEEDSPTKKSKSLDTSIFQADQFRHPVSSKLEN 180

Qy 68 LD-----OKTSHLTDELDALFDVQKARAVRGLVNEEDSD----- 105

Db 181 SELSEVDKPPFTASRRSRKRPVSAESDE-DEDFDAPTKGSRHKRIVSDSDSDVPEPHI 239

Qy 106 ----SSQSFVRDKYMSNGTIDNFSLDTAITMPTGTPRSDDGDLFFGDKSKODASNDV 161

Db 240 SEASSEASLPIDEVSMDEDDVGYSDHSVSAAPIPK-----KSRKSSNSL 287

Qy 162 VEELR 166

Db 288 YESYR 292

# RESULT 5

T49790  
 related to TOM1 protein [imported] - Neurospora crassa  
 N:Alternate names: protein B11B22.10  
 C:Species: Neurospora crassa  
 C:Date: 02-Jun-1979 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
 R:Schneider, U.; King, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu  
 submitted to the Protein Sequence Database, May 2000  
 A:Reference number: Z25022  
 A:Accession: T49790  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-3839 <SCH>  
 A:Cross-references: EMBL:AL356834; GSPDB:GN001116; NCSP:B11B22.10  
 A:Experimental source: BAC clone B11B22; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B11B22.10  
 A:Map position: 6  
 A:Introns: 16/3; 2607/1; 2623/1; 2658/1; 2845/1; 2987/2; 3204/3; 3694/1; 3809/1

Query Match 9.8%; Score 91.5; DB 2; Length 3839;  
 Best Local Similarity 24.6%; Pred. No. 50;  
 Matches 47; Conservative 29; Mismatches 72; Indels 43; Gaps 9;



A:Title: Cloning of the Plasmodium vivax Duffy receptor.

A:Reference number: 220901; MUID:91187056

A:Accession: T30848

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1070 <FAN>

A:Cross-references: EMBL:M61095; NID:g160275; PID:g160276; PIDN:AAA63423.1

C:Genetics:

A:Introns: 19/3; 1006/1; 1032/2; 1057/1

A:Note: PVDR

Query Match 9.5%; Score 89.5; DB 2; Length 1070;  
Best Local Similarity 24.0%; Pred. No. 16;  
Matches 44; Conservative 30; Mismatches 62; Indels 47; Gaps 8;

Oy 6 LRHSVNSQLSGSPAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAAASAI 65

Db 671 LRHSKNSDSGPA-----ESMNPDSNSKGETG-----KGQNDMARATWDSSNS 717

Oy 66 LGLDQKSHLTDELDALFDVOKARVRGLVEDNEDSDSSQSFVPRDKYMSN---GTI 122

Db 718 DG-----TSATGTTDAV-----DREINKGVPE-----RDKTGSKDGGGE 755

Oy 123 DNFSLDTAITMPTGPRSDGDLFFGDKKSKODASNDVVEE---LRQOQOAMEEALKTA 179

Db 756 DNSANKDAATVGEDRIENSAG-----GSTNDRSKNTEKNGASTPDSKQSEDATALS 809

Oy 180 KOE 182

Db 810 KTE 812

RESULT 10

T36595

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-476 <FAN>

A:Cross-references: EMBL:AL109987; PIDN:CA83399.1; GSPDB:GN00070; SCODEB:SCF20.01c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SCF20.01c

Query Match 9.5%; Score 89; DB 2; Length 476;

Best Local Similarity 36.2%; Pred. No. 6.5;

Matches 21; Conservative 7; Mismatches 30; Indels 0; Gaps 0;

Oy 5 LLHSVNSQLSGSPAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAA 62

Db 268 LLEAFTDRPADAPGPGTSGRTGTHSGRAEGAAGRPPLPCLAGTDPANWSYSRTAA 325

RESULT 11

T00347

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-476 <FAN>

A:Cross-references: EMBL:AL109987; PIDN:CA83399.1; GSPDB:GN00070; SCODEB:SCF20.01c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SCF20.01c

Query Match 9.5%; Score 89; DB 2; Length 476;

Best Local Similarity 36.2%; Pred. No. 6.5;

Matches 21; Conservative 7; Mismatches 30; Indels 0; Gaps 0;

Oy 5 LLHSVNSQLSGSPAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAA 62

Db 268 LLEAFTDRPADAPGPGTSGRTGTHSGRAEGAAGRPPLPCLAGTDPANWSYSRTAA 325

RESULT 11

T00347

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-476 <FAN>

A:Cross-references: EMBL:AL109987; PIDN:CA83399.1; GSPDB:GN00070; SCODEB:SCF20.01c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SCF20.01c

Query Match 9.5%; Score 89; DB 2; Length 476;

Best Local Similarity 36.2%; Pred. No. 6.5;

Matches 21; Conservative 7; Mismatches 30; Indels 0; Gaps 0;

Oy 5 LLHSVNSQLSGSPAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAA 62

Db 268 LLEAFTDRPADAPGPGTSGRTGTHSGRAEGAAGRPPLPCLAGTDPANWSYSRTAA 325

RESULT 11

T00347

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-476 <FAN>

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1353 <NAG>

A:Cross-references: EMBL:AB011164; NID:g3043707; PIDN:BAA25518.1; PID:g3043708

A:Experimental source: brain; clone HJ2807

C:Genetics:

A:Note: KIAA0592

Query Match 9.4%; Score 88.5; DB 2; Length 1353;

Best Local Similarity 24.4%; Pred. No. 25;

Matches 50; Conservative 19; Mismatches 61; Indels 75; Gaps 9;

Oy 13 SOLGSPAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAAAILGLDQKI 72

Db 474 SPYGPPTGLFD-----DDGDD-----DDDFSPHSPKSTRKV-QST 512

Oy 73 SHLTDDLDLAF-----DDVOKARA-----VRRGLVE 99

Db 513 ADIFGDEBGLFKEKAVASPEATVSDENKARAKVKVTLSSSKNLKPSSETKTKGLFS 572

Oy 100 DNEDSD-----SQSSFPVRDKYSNGTIDNFSLDTAITMPTGPRSDGDLFFGDKKSKQ 155

Db 573 DEEDSEDLSQSASNLKGASLLPG-----KLPTSVSLFD---DEEDENLFGTGAARKQ 624

Oy 156 DASNVDEELRQOQAMEEALKTA 180

Db 625 TLS-----LQAGREKAKASE 640

RESULT 12

S51796

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-384 <HAF>

A:Cross-references: EMBL:Z46388; NID:g624249; PIDN:CAA86522.1; PID:g624250

A:Experimental source: protein G7056; protein YGR167W

C:Keywords: phosphoprotein

Query Match 9.4%; Score 88; DB 2; Length 384;

Best Local Similarity 25.6%; Pred. No. 6.1;

Matches 52; Conservative 26; Mismatches 75; Indels 50; Gaps 9;

Oy 18 PAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAAAILGLDQKSHLTD 77

Db 187 PPGVSLSGSAGHAGGAGGPPAPPPLPQTSGGTGAPGLAA-ATAG--AKLRKYSK 243

Oy 78 DELDALFDVOKARVR---GLVED-----NEDSDSQSFSF 110

Db 244 QEASGCGPPVPRKASTSTGGLMEENAMLARRKATQVGEKPKDKSANEPEARVPV 303

Oy 111 P-----VRDKYMSNGTI-----DNFSLDTAITMPTGPRSDGDLFFGDKKSKQDASNV 160

Db 304 PAQSETVRRPWKXNSTTLPKMKSSSVTSEAHFSTPSSDSDDL-----ERYKQEL--- 355

Oy 161 DVEELRQO-QAOOMEALKTAQOE 182

Db 356 -LEEYRKELQVKEIIEAFVQE 377

RESULT 13

A36425

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-384 <HAF>

A:Cross-references: EMBL:Z46388; NID:g624249; PIDN:CAA86522.1; PID:g624250

A:Experimental source: protein G7056; protein YGR167W

C:Keywords: phosphoprotein

Query Match 9.4%; Score 88; DB 2; Length 384;

Best Local Similarity 25.6%; Pred. No. 6.1;

Matches 52; Conservative 26; Mismatches 75; Indels 50; Gaps 9;

Oy 18 PAGRHRTLSKAGSDAGDSRSDPLPHFRGKSDMFYARREAAAILGLDQKSHLTD 77

Db 187 PPGVSLSGSAGHAGGAGGPPAPPPLPQTSGGTGAPGLAA-ATAG--AKLRKYSK 243

Oy 78 DELDALFDVOKARVR---GLVED-----NEDSDSQSFSF 110

Db 244 QEASGCGPPVPRKASTSTGGLMEENAMLARRKATQVGEKPKDKSANEPEARVPV 303

Oy 111 P-----VRDKYMSNGTI-----DNFSLDTAITMPTGPRSDGDLFFGDKKSKQDASNV 160

Db 304 PAQSETVRRPWKXNSTTLPKMKSSSVTSEAHFSTPSSDSDDL-----ERYKQEL--- 355

Oy 161 DVEELRQO-QAOOMEALKTAQOE 182

Db 356 -LEEYRKELQVKEIIEAFVQE 377

RESULT 13

A36425

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-384 <HAF>

A:Cross-references: EMBL:Z46388; NID:g624249; PIDN:CAA86522.1; PID:g624250

A:Experimental source: protein G7056; protein YGR167W

C:Keywords: phosphoprotein

Query Match 9.4%; Score 88; DB 2; Length 384;

Search completed: April 25, 2001, 10:14:18  
Job time: 294 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2001, 10:18:33 ; Search time 36.31 Seconds  
(without alignments)  
238.265 Million cell updates/sec

Title: US-09-235-416-1\_COPY\_602\_784

Perfect score: 938  
Sequence: 1 OPOSLLRHVSNTQLSGSPAP.....ELRQQQAQWEALTKAQEF 183

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106.5	11.4	376	1 VASP_MOUSE	P70460 mus musculo
2	91.5	9.8	841	1 DES3_MCHWS	Q95194 murine cyto
3	89.5	9.5	1070	1 VASP_PLAYS	P22290 plasmodium
4	88.5	9.3	1084	1 VASP_PLAYS	P22290 plasmodium
5	87.5	9.3	233	1 CLC1_YEAST	P17891 canis famil
6	86.5	9.2	1957	1 YDR6_YEAST	P17891 schizosacch
7	85.5	9.1	629	1 RPSD_HAFTN	Q10411 schizosacch
8	83.5	8.9	1203	1 XCPD_XENLA	P43756 haemophilus
9	82.5	8.8	685	1 SNWA_DICDI	P50533 xenopus lae
10	82	8.7	598	1 YAA5_SCHPO	P54705 dicystostell
11	82	8.7	706	1 Z1S1_CHICK	Q09799 schizosacch
12	82	8.7	1130	1 YL17_CHAEL	Q06235 gallus gall
13	82	8.7	2131	1 CINA_DROME	Q11102 caenorhabdi
14	81.5	8.7	638	1 Y053_HUMAN	P35500 drosophila
15	81.5	8.7	655	1 HS70_DAUCA	P42331 homo sapien
16	81.5	8.7	809	1 PUPB_PSEPU	P26791 daucus caro
17	81.5	8.7	852	1 CSG_HAHLA	P38047 pseudomonas
18	81.5	8.7	1053	1 SPFS_YEAST	P08198 halobacteri
19	81	8.6	772	1 RA15_YEAST	P27692 saccharomyc
20	80.5	8.6	101	1 KPSA_MOUSE	P46599 schizosacch
21	80	8.5	1095	1 TNP5_CHAEL	P28741 mus musculu
22	80	8.5	2349	1 TNP5_YEAST	Q03553 caenorhabdi
23	79.5	8.5	1406	1 SVB9_YEAST	P12593 homo sapien
24	79.5	8.5	1347	1 YB77_YEAST	P35394 saccharomyc
25	78.5	8.4	1381	1 DPOA_YEAST	P34216 saccharomyc
26	78.5	8.4	1468	1 YDPT_YEAST	P13392 saccharomyc
27	78	8.3	204	1 HS2C_CHERU	P11890 chenopodisp
28	78	8.3	219	1 AG1X_ARATH	Q02743 arabidopsis
29	78	8.3	379	1 KAPO_BOVIN	P00514 bos taurus
30	78	8.3	489	1 DMP1_RAT	P98193 rattus norv
31	78	8.3	681	1 MP10_HUMAN	O00566 homo sapien
32	78	8.3	960	1 UVPA_TREPA	O83527 treponema p
33	78	8.3	1938	1 MYS_AEQIR	P24733 aequipecten

34	77.5	8.3	660	1 HT31_ARATH	O04596 arabidopsis
35	77.5	8.3	1022	1 DPOS_YEAST	P39585 saccharomyc
36	77.5	8.3	1025	1 MK21_YEAST	Q12176 saccharomyc
37	77.5	8.3	1658	1 YM67_YEAST	Q03661 saccharomyc
38	77	8.2	439	1 MFAL_HUMAN	P55081 homo sapien
39	77	8.2	881	1 YIT8_YEAST	P39535 saccharomyc
40	77	8.2	1526	1 TP2A_RAT	P41516 rattus norv
41	76.5	8.2	357	1 FRBA_DROME	P54397 drosophila
42	76.5	8.2	357	1 MSP1_CAEEL	P54815 caenorhabdi
43	76.5	8.2	618	1 NOP2_YEAST	P40991 saccharomyc
44	76.5	8.2	667	1 R1P4_TOBAC	Q03684 nicotiana t
45	76.5	8.2	821	1 YNZO_CAEEL	P45970 caenorhabdi

#### ALIGNMENTS

RESULT	1
VASP_MOUSE	
ID	VASP_MOUSE
AC	P70460; Q9R214;
DT	01-OCT-2000 (Rel. 40, Created)
DT	01-OCT-2000 (Rel. 40, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	VASODILATOR-STIMULATED PHOSPHOPROTEIN (VASP).
GN	VASP.
OS	Mus musculus (Mouse)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eumammalia; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=129;
RX	MEDLINE=96411679; PubMed=8812448;
RA	Zimmer M., Fink T., Fischer L., Hauser W., Scherer K., Lichter P.,
RA	Walter U.;
RT	"Cloning of the VASP (vasodilator-stimulated phosphoprotein) genes in
RT	human and mouse: structure, sequence, and chromosomal localization.";
RL	Genomics 36:227-233(1996).
RN	[2]
RP	SEQUENCE FROM N.A., AND PHOSPHORYLATION.
RC	STRAIN=C57BL/6J;
RX	MEDLINE=99185054; PubMed=10085070;
RA	Collins S.P., Uhler M.D.;
RA	Cyclic AMP- and cyclic GMP-dependent protein kinases differ in their
RT	regulation of cyclic AMP response element-dependent gene
RT	transcription.";
RL	J. Biol. Chem. 274:8391-8404(1999).
CC	-1- FUNCTION: ACTIN- AND PROFILIN-BINDING MICROFILAMENT-ASSOCIATED
CC	PROTEIN. MAY ACT IN CONCERT WITH PROFILIN TO CONVEY SIGNAL
CC	TRANSDUCTION TO ACTIN FILAMENT PRODUCTION (BY SIMILARITY).
CC	-1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: FOCAL ADHESIONS (BY SIMILARITY).
CC	-1- PTM: MAJOR SUBSTRATE FOR CAMP-DEPENDENT (CAPK) AND CGMP-DEPENDENT
CC	PROTEIN KINASE (CGPK) IN PLATELETS.
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by, and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
DR	EMBL; X98475; CAA67108.1;
DR	EMBL; AF084548; AA016045.1;
DR	MGD; MG1:109268; VASP.
DR	InterPro; IPR000697;
DR	InterPro; IPR001960;
DR	Pfam; PF00568; WH1.1.
KW	Phosphorylation; Actin-binding.
FT	DOMAIN 166 182 POLY-PRO.
FT	DOMAIN 318 321 POLY-SER.

	Query Match	9.5%	Score 89.5	DB 1	Length 1070
	Percent Identity	24.0%	Percent ID 9.6		
	Matches	44	Conservative	30	Mismatches 62; Indels 47; Gaps
KW	Meloidin, Receptor, 21				
FT	SIGNAL	1			POTENTIAL .
FT	CHAIN	21	1070		DUFFY RECEPTOR .
FT	DOMAIN	21	1007		POTENTIAL .
FT	TRANSMEM	1008	1025		EXTRACELLULAR (POTENTIAL) .
FT	DOMAIN	1026	1070		POTENTIAL .
FT	CARBOHYD	183	183		CYTOSOLASMIC (POTENTIAL) .
FT	CARBOHYD	255	255		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	255	255		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	351	351		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	451	451		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	715	715		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	787	787		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	825	825		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	903	903		N-LINKED (GLCNAC. .) (POTENTIAL) .
FT	CARBOHYD	938	938		N-LINKED (GLCNAC. .) (POTENTIAL) .
SQ	SEQUENCE	1070 AA	119683 MW	CB051DF13E294603	CRG64

Db 671 LRHSKNSDSGPA-----ESMANPDSNKGCTG---KGQNDMAKATKDSNSS 717

Qy 66 LGLDQKISHLTDELDALFDVDQKARVRRGLVEDNEDSDSQSSPVRDKYMSN---GTI 122

```

Db 718 DG-----TSSGDTTDAV-----DREINKGVPED-----RDKTVGSKDGGGE 755
Qy 123 DNFSLDTAITWPTPRSDGDLFGDKSKQDASNVNVEE-----LRQQAQMEERALKTA 179
Db 756 DNSAKDAATVVGEDRIENSAG-----GSTNORSKNDTEKNGASTPQSKQSDATKALS 809
Qy 180 KQE 182
Db 810 KTE 812

RESULT 4
VASP_CANFA STANDARD; PRT; 384 AA.
AC P50551;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE VASODILATOR-STIMULATED PHOSPHOPROTEIN (VASP).
GN VASP.
OS Canis familiaris (Dog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
EN 111
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95125547; PubMed=7828592;
RA Hafner C., Jarchau T., Reinhard M., Hoppe J., Lohmann S.M.,
RA Walter U.;
RT *Molecular cloning, structural analysis and functional expression of
RT the proline-rich focal adhesion and microfilament-associated protein
RT VASP.*;
RL EMBJ J. 14:19-27(1995).
CC -1- FUNCTION: ACTIN- AND PROFILIN-BINDING MICROFILAMENT-ASSOCIATED
CC PROTEIN. MAY ACT IN CONCERT WITH PROFILIN TO CONVEY SIGNAL
CC FROM PRODUCING CELL TO ACTIN FILAMENT PRODUCTION.
CC -1- SUBCELLULAR LOCATION: FOCAL ADHESIONS.
CC -1- PTM: MAJOR SUBSTRATE FOR CAMP-DEPENDENT (CAPK) AND GMP-DEPENDENT
CC PROTEIN KINASE (CGPK) IN PLATELETS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; 246388; CAA8522.1; -.
CC InterPro: IPR001900; -.
CC Pfam: PF00568; WH1; 1.
CC Phosphorylation: Actin-binding.
CC DOMAIN 173 184 POLY-PRO.
CC DOMAIN 173 184 POLY-PRO.
CC DOMAIN 326 329 POLY-SER.
CC FT MOD_RES 160 160 PHOSPHORYLATION (BY CAPK AND CGPK)
CC (BY SIMILARITY).
CC FT MOD_RES 242 242 PHOSPHORYLATION (BY CAPK AND CGPK)
CC (BY SIMILARITY).
CC FT MOD_RES 281 281 PHOSPHORYLATION (BY CAPK AND CGPK)
CC (BY SIMILARITY).
CC FT MOD_RES 281 281 PHOSPHORYLATION (BY CAPK AND CGPK)
CC (BY SIMILARITY).
CC SEQUENCE 384 AA: 40413 MW; 2668B3C46FB4397F CRC64;
Query Match 9.4%; Score 88; DB 1; Length 384;
Best Local Similarity 25.6%; Pred. No. 3.7;
Matches 52; Conservative 26; Mismatches 75; Indels 50; Gaps 9;
Qy 18 PAPCRHDTLTAKGSDADGDSPLPHEPKGSDWYFARRAASAIGLDQKISHUTD 77
Db 187 PPGVSLSGSGANGAGGGPPAPLPPLTAQTSGGTGAGGLAA-ALAG--AKLRKYSK 243

```

```

Qy 78 DELDALPDVOKARAVRR-----GLVED-----NEDSDSOSF 110
Db 244 QEASGGFPVKRASTGTGGMEENMLARRKATQGVKKPKDSEAEFEAPVP 303
Qy 111 P-----VRKYNSNGTI-----DNFSLDTAITWPTPRSDGDLFGDKSKQDASNV 160
Db 304 PAQSETVRPRAEINSTTLPRMKSSSVTTSENPSTSSDSDSL-----ERVQEL---- 355
Qy 161 DVEELROO-QAGNEEALAKQOE 182
Db 356 -LEEVKELQKVEELIEAFVQE 377

RESULT 5
CLCL_YEAST STANDARD; PRT; 233 AA.
AC P17891;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE CLATHRIN LIGHT CHAIN (CLC).
GN CLC1 OR CLC16W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Eungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
EN 111
RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-39.
RX MEDLINE=91009480; PubMed=2211819;
RA Silveira L.A., Mong D.H., Maslarz F.R., Schekman R.;
RT *Yeast clathrin has a distinctive light chain that is important for
RT cell growth.*;
RL J. Cell Biol. 111:1437-1449(1990).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITTS AND
CC HEAVY CHAINS AND 3 LIGHT CHAINS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITTS AND
CC HEAVY CHAINS AND 3 LIGHT CHAINS.
CC -1- MYOGLIUM.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X52272; CAA36515.1; -.
CC EMBL; 272952; CAA97192.1; -.
CC EMBL; 272953; CAA97193.1; -.
CC FIC: A36425; A36425.
CC SGD: S0003399; CLCL.
CC InterPro: IPR000936; -.

```

```

DR Pfam; PF01086; Clathrin_lg_ch.1.
DR PROSITE; PS00244; CLATHRIN_LIGHT_CHN.1; FALSE_NEG.
DR PROSITE; PS00581; CLATHRIN_LIGHT_CHN.2; 1.
KW Coated pits; Calcium-binding; Calmodulin-binding; Coiled coil.
FT BINDING 144 204 HEAVY CHAIN (POTENTIAL).
FT SIMILAR 144 204 TO REGIONS OF INTERMEDIATE FILAMENTS AND
FT OTHER PROTEINS THAT FORM COILED-COIL
FT STRUCTURES.
FT SIMILAR 193 213 TO THE BRAIN SPECIFIC INSERT IN LCA/LCB.
SQ SEQUENCE 233 AA; 26531 MW; 27BAB175780ECB3 CRC64;

Query Match 9.3%; Score 87.5; DB 1; Length 233;
Best Local Similarity 27.6%; Pred. No. 2.2;
Matches 40; Conservative 22; Mismatches 54; Indels 29; Gaps 8;

QY 50 KDSDFYARAEASAIIGLQKISHLTDDLEDAFDVOKAVRGLVDENEDSDSS 109
DB ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
21 KDDDTDFLKREA--EILG-----DEFTQDDILETEA--SPAKDDDETRFEQ 66

QY 110 FPRVRYKMSNGTIDNFSIDTAITMPTSRDDGQ-ALFFGDKKS-----KQDASNDVVEE 164
DB ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
67 FP--DINSANGAVSSDNGSATVSSGNDNGEADDDFSTFEGANSTESVKDRSEV-VDQ 123

QY 165 LROOQA-----QMEALKATAKOE 182
DB ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
124 WKQRAVEIHKDLKDELKELQD 148

RESULT 6
YD86_SCHPO STANDARD: PRT: 1957 AA.
AC Q10411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 222.8 KDA PROTEIN C1F3.06C IN CHROMOSOME I.
GN SPAC1F3.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX Schizosaccharomycetes.
NX NCBI_Taxid=4896;
RA [1]
RA SEQUENCE FROM N.A.
RC STRAIN=972;
RC Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 270690; CAA94624.1;
CC Hypothetical protein.
CC SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

Query Match 9.2%; Score 86; DB 1; Length 1957;
Best Local Similarity 23.1%; Pred. No. 38;
Matches 45; Conservative 41; Mismatches 65; Indels 44; Gaps 10;

QY 1 QEO--SLLRHSVTSNQLSGSPAGRHDTLTSGKSDAGDSPLPHFRKGSDFYAR 58
DB 954 QEEISNLKNEENSQOATTSVKSLDETLKS-----SKLEADIEHLKNKYSEVEVER 1006
QY 59 REAASAIIGLQKISHLTDD-----ELDALFDVOKAVRGLVDENEDSDSSFPV 112
DB ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1007 ----NALLASNER--LMDDLKNNCNIASLTQTELEKRA-----ENDLQSLSV-V 1051

```

```

QY 113 RDVYMSNGTIDNFSIDTAITMPTSRDDGQALFFGDKK-----SKQDASNDVVEELRQ 167
DB ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1052 SSEY-----ENLL--ISSQNSLEDTQWLYKRNKWLKDEKQORVLEULTS 1102

QY 168 QQAQMEALKATAKOE 182
DB ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
1103 KYKGLGEENAIKIDE 1117

RESULT 7
RPSD_HAEIN STANDARD: PRT: 629 AA.
AC P43766;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RNA POLYMERASE SIGMA FACTOR RPOD (SIGMA-70).
GN RPOD OR HT0533.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RA [1]
RA SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Influenzae Rd.;
CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC BACTERIA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32735; AAC22190.1;
CC HSSP; P00579; IS1G.
CC TIGR; H10533;
CC InterPro; IPR000943;
CC Pfam; PF00140; sigma70; 1.
CC PRINTS; PR00046; SIGMA70ECT.
CC PROSITE; PS00715; SIGMA70.1;
CC PROSITE; PS00716; SIGMA70.2; 1.
CC Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
CC DNA-binding. 119 432 POLYMERASE CORE BINDING (POTENTIAL).
KW DOMAIN BIND 587 602 H-T-H MOTIF (BY SIMILARITY).
FT DOMAIN BIND 587 602
SQ SEQUENCE 629 AA; 72084 MW; FCC64EB8CBFB9CF5 CRC64;

Query Match 9.1%; Score 85.5; DB 1; Length 629;
Best Local Similarity 21.8%; Pred. No. 11;
Matches 44; Conservative 31; Mismatches 88; Indels 39; Gaps 7;

QY 3 QSLLRHSVTSNQLSGSPA-FGRHDTLTSGKSDA-----DQDSRSDSPLPHFRKGSDFW 55

```

Db 26 OCYLTAETINLLPDAIPEDYVYKLTQNDAGIPVLDAPESDQWMLSDTFPED-- 83  
QY 56 YAREAAASAILGDKQKLSHTLDD-----ELDALFDVOKARAVRGLVEDNES 104  
Db 84 -AVEATEOLLNSVEISGTITDPVRYVREMGTVDLLTREDISTAKREGGI--DEVQ 139  
QY 105 DSQSSFPVRKYSNGTIDNF-----SLDTATMPGTSPRSDDGDLFFGDKKS 153  
Db 140 TSIATP-----EALNGLKANTDOVEKGNFRJDLITGVPDPAETEEHNGDSEDD 195  
QY 154 KQASWVDEELRQQAQWEEA 175  
Db 196 BESSNAVDNEDEDENES 217  
RESULT 8  
XCPE\_XENLA STANDARD; PRT: 1203 AA.  
ID XCPE\_XENLA AC P50533;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 35, Last annotation update)  
DE CHROMOSOME ASSEMBLY PROTEIN XCAP-E.  
GN XCAP-E.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
OX Xenopodidae; Xenopus.  
CA NCBI\_TaxID=8355;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=55042742; PubMed=7954011;  
RA Hirano T, Mitchison T J.  
RT "A heterodimeric coiled-coil protein required for mitotic chromosome  
condensation in vitro";  
RL Cll 79:449-458(1994).  
CC -1- FUNCTION: REQUIRED FOR BOTH ASSEMBLY AND STRUCTURAL MAINTENANCE OF  
MITOTIC CHROMOSOMES.  
CC -1- SUBUNIT: ASSOCIATES WITH XCAP-C PROBABLY AS HETERODIMER.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR. DURING CHROMOSOME ASSEMBLY IN  
MITOTIC EXTRACTS, XCAP-C/E WAS RECRUITED TO THE CHROMATIN AND  
FORMED A DISCRETE INTERNAL STRUCTURE WITHIN ASSEMBLED CHROMOSOMES.  
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS.  
CC -1- FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.  
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: U13674; AAA64680.1; .  
DR HSPSP; P07751; IAU3.  
KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.  
FT NP\_BIND 32 39 ATP (POTENTIAL).  
FT DOMAIN 172 356 COILED COIL (POTENTIAL).  
FT DOMAIN 397 513 COILED COIL (POTENTIAL).  
FT DOMAIN 670 1032 COILED COIL (POTENTIAL).  
FT DOMAIN 1095 1121 ALA/ASP-RICH (DA-BOX).  
SQ SEQUENCE 1203 AA; 136341 MW; 04323DD0027DF309 CRC64;

Query Match 8.98; Score 83.5; DB 1; Length 1203;  
Best Local Similarity 20.34; Pred. No. 33;  
Matches 29; Conservative 29; Mismatches 68; Indels 17; Gaps 3;  
QY 47 FRKGSDNFYAREAAATGCLGDKQKLSHTLDD-----ELDALFDVOKARAVRGLVEDNES 106  
Db 713 YROLKQWEMKEAEALLQTKLQSSVYHQEELDSLQATIESEETLANKTEVQKAAEE 772

QY 107 OSSPFRDYKYSNGTID-----NFSLDTATMPGTSPRSDDGDLFFGDKSKODAS 158  
Db 773 K--PVVLEHKMKNAEAREERELKEAQKLDTA-----KKDAASNKKMKKEQKQVFDAL 823  
QY 159 NVDEELRQQAQWEEALKTKAKO 181  
Db 824 VLEELKEQTYTKQIETVDE 846  
RESULT 9  
SNWA\_DICDI STANDARD; PRT: 685 AA.  
ID SNWA\_DICDI AC P54705;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 35, Last annotation update)  
DE SNWA PROTEIN.  
GN SNWA  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=14689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97128797; PubMed=8973337;  
RA Folk P, Puta F, Krpejsova L, Blahuskova A., Markos A.,  
RA Rabino M., Dottin R.P.;  
RT "The homolog of chromatin binding protein Bx42 identified in  
dictyostelium";  
RL Gene 181:229-231(1996).  
CC -1- SIMILARITY: BELONGS TO THE SNW FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: U43887; AAA0497.1; .  
DR Dictyob; DD00074; snwa.  
FT DOMAIN 31 41 POLY-GLN.  
FT DOMAIN 194 360 SNW DOMAIN.  
FT DOMAIN 245 253 PRO-RICH.  
FT DOMAIN 409 415 POLY-ASP.  
FT DOMAIN 539 616 SH2-LIKE DOMAIN.  
SQ SEQUENCE 685 AA; 78529 MW; 1DC9521E997A593 CRC64;  
Query Match 8.98; Score 82.5; DB 1; Length 685;  
Best Local Similarity 24.94; Pred. No. 20;  
Matches 48; Conservative 26; Mismatches 86; Indels 31; Gaps 8;  
QY 1 QEQSLR--HSVTNSQLSGSPAGCHDHTLSKAGSDAGDSKRSPLPHPRGKSDWFA 57  
Db 339 RKQDMKLKLAEDVNRERSGI----IQQVYTRKDNSSDNDNDNDSSDDKDKKTRPPNRR 394  
QY 58 RREAAASAILGDKQKLSHTLDD-----ELDALFDVOKARAVRGLVEDNES-----DSQSSF 110  
Db 395 RRSRSY-----ERIPSRNDND--DDRYIKDNDRGNRDNIDSRNDRSRDRSR 446  
QY 111 PYRDYKYSNGTIDNF-----SLDTATMPG--TPRSDDGDLFF-----GDKKSKQDASND 161  
Db 447 DSRDSRDRSRDRSRDRSRDRSRDRSRDRSRDRSRDRSRDRSRDRSRDRSRDRSRDRSR 506  
QY 162 VEELRQQAQWEE 174  
Db 507 DENVRRERKELE 519  
RESULT 10  
YAA5\_SCHPO

RX

Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

1- SIMILARITY: WEAK, TO MYOINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL 041545: NA893190.1.  
 Accession: CO2F12.7, CE031501.  
 Hypothetical protein: Coiled coil.  
 DOMAIN 121 779 COILED COIL (POTENTIAL).  
 DOMAIN 805 1061 COILED COIL (POTENTIAL).  
 SEQUENCE 1130 AA: I31485 MW: 80FD2EEFD99FB09 CRC64:

Query Match 8.7%: Score 82; DB 1; Length 1130;  
 Best Local Similarity 21.2%; Pred. No. 41;  
 Matches 40; Conservative 34; Mismatches 55; Indels 60; Gaps 9;

QY 27 LKSGDAGDSQSRSPPLPHF---GKSDWYFYAREASALGLQKQIKSHLTDEIDA 82  
 DD 566 LAKIVETSQUTRI---LEETREITGDAELILNRQK---LEKELSH-TERDEL 614  
 QY 83 LFEDVQK---ARARV---RGVDESDSDSSFPDVKYASGTT-- 122  
 DB 615 LHNTQKELEAKHKTHTVTVLRGAEIDQKSAFENQRYGKESAKTRLEAQNKTLIS 674  
 QY 123 ---DNFSLDTATMPTGTRSDDDGDALEFGDKKSQAQSDVNDVEELRQQAQME 173  
 DB 675 EMKVKIVAEMLAETS-----DKDNLL-----EELSKNKNIHLKQETAGLN 718

174 EALKTAQKE 182  
 DB 719 EKISTKTE 727

RESULT 13  
 CINA DROME  
 ID CINA DROME STANDARD; PRT: 2131 AA.

AC P35500:  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 GN SODIUM CHANNEL PROTEIN PARA (PARALYTIC PROTEIN).  
 GN PARA.  
 GN Drosophila melanogaster (Fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Phlebotominae; Simuliidae; Drosophila.  
 NCBI\_TAXID:727,  
 [1]  
 RP SEQUENCE OF 1-1862 FROM N.A.  
 RP STRAIN-CANTON-S; TISSUE-Head.  
 RC MEDLINE-89376565; PubMed-2550145;  
 RC Loughney K., Kreber R., Ganetzky B.;  
 RT "Molecular analysis of the para locus, a sodium channel gene in  
 RT Drosophila".  
 RT Cell 58:1143-1154(1989).  
 RP [2]  
 RP SEQUENCE OF 1693-1895 FROM N.A.  
 RP MEDLINE-89184571; PubMed-2538830;  
 RA Ramawami M., Tanouye M.A.;  
 RA "Two sodium-channel genes in Drosophila: implications for channel  
 RP distribution".  
 RP Proc. Natl. Acad. Sci. U. S. A. 86:2079-2082(1989).

1- FUNCTION: THIS PROTEIN MEDIATES THE VOLTAGE-DEPENDENT SODIUM ION PERMEABILITY OF EXCITABLE MEMBRANES. ASSUMING OPENED OR CLOSED CONFORMATIONS IN RESPONSE TO THE VOLTAGE DIFFERENCE ACROSS THE MEMBRANE, THE PROTEIN FORMS A SODIUM-SELECTIVE CHANNEL THROUGH

```

FT VARIANT 189 189 Y -> C (IN 2520.2).
FT VARIANT 198 198 M -> V (IN 2520.2).
FT VARIANT 205 205 C -> R (IN 2520.2).
SQ SEQUENCE 2131 AA: 239385 MW: DLECE68845A90F1 CRC64;

Query Match
Best Local Similarity 22.4%; Pred. No. 88;
Matches 43; Conservative 31; Mismatches 60; Indels 58; Gaps 9;

Oy 5 LLHSVTSNLSGSPAGPGRHRTLSKAGSDAGDSRSDPLHFRGKDSWFYARREAASA 64
Db 1084 LIRNKLTVN-QISDSQSGERNTOISWISBKGVCV-----CISA 1121
Oy 65 ILGLDQKISHLTDELDALFDVQKARAVRGLVEDNE-----DSQSOSFFP 111
Db 1122 EHG-----DNEILGHDELADSLIKKGIKEQVLEVAIGDQWETHGDMKNKP 1172
Oy 112 VRKYNMNGTIDNFSLDTA-ITMPTGPRSDDDGALFFG-----DKKSQDASNDV 162
Db 1173 KKSGLNNATDD-----DTASINSYGHKNRPFKOESHKSGAETMEGEKRDASKEDLGLD 1228
Oy 163 EELRQOQAE 174
Db 1229 EEL-DEEGEE 1239

RESULT 14
Y053_HUMAN
ID Y053_HUMAN STANDARD: PRT: 638 AA.
AC P42331;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHELICAL PROTEIN KIAA0053.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051396; Pubmed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki T., Kawarabayashi T., Ishikawa K.-I., Tabata S.;
RA Prediction of the coding sequences of unidentified human genes. II.
RA The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RA analysis of complementary DNAs from human cell line KG-1.;
RA DNA RES. 223-225(1994).
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z96642; BAA06125.1; -.
CC HSP: Q07960; IBCG.
CC InterPro: IPR000198; -.
CC InterPro: IPR001849; -.
CC Pfam: PF00169; PH_1.
CC PROSITE: PS00620; R06GAP.1.
CC PROSITE: PS50003; PH DOMAIN: 1.
CC Hypothetical protein.
CC DOMAIN 39 144 PH
CC SEQUENCE 638 AA: 72431 MW: 082897EB4A880871 CRC64;

Query Match
Best Local Similarity 8.7%; Score 81.5; DB 1; Length 638;
Matches 3; Conservative 21; Mismatches 40; Indels 29; Gaps 5;

Oy 67 GLDQKIS-----HLTDELDALFDVQKARAVRGLVEDNDSQSOSFFPRKYMSNG 120
Db 504 GVANKLITTDNGRLSKDIEKLVEKYKA-----EED-----VKKYKAKN 548
Oy 121 TIDNFSLDTAITMPTGPRSDDDGALFFGDKSKQDASNDVDEELRQQA-----OMEAL 176

```



Mon Apr 30 07:47:12 2001

us-09-235-416-1\_copy\_602\_784.rsp

Page 9

Db 549 ALENY-----AYNNRNTIKDDKIFGKLDAGDKEKIEIETAVNEAIEWLEKNQLAEVEDEKDL 604

Qy 177 K 177

Db 605 K 605

Search completed: April 25, 2001, 10:18:38  
Job time: 304 sec

**THIS PAGE BLANK (USPTO)**

ON protein - protein search, using sw model  
 Run on: April 25, 2001, 10:18:12 ; Search time 77: 5 Seconds  
 without alignments  
 276,762 Million cell updates/sec

## ALIGNMENTS

RESULT	1
Q9R214	
ID	Q9R214 PRELIMINARY; PRT; 375 AA.
AC	Q9R214;
DT	01-MAY-2000 (TREMBLrel_13, Created)
DT	01-MAY-2000 (TREMBLrel_13, Last sequence update)
DT	01-OCT-2000 (TREMBLrel_15, Last annotation update)
DE	VASODILATOR-STIMULATED PHOSPHOPROTEIN.
DE	VASP.
GN	Mus musculus (Mouse).
OS	Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Metazoa; Rodentia; Sciurognathi; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
SEQUENCE FROM N.A.	
KC	STRAIN=C57BL/6J;
KC	Collins S.P., Uhler M.D.;
RA	Cyclic AMP- and Cyclic GMP-dependent Protein Kinases Differ in Their
RA	Regulation of Cyclic AMP Response Element-dependent Gene
RT	Transcription on: 374-8391-8404 (1999).
RT	J Biol Chem 274(18):13400-13406 (1999).
DR	EMBL: AF084548; ADI15045.1; -.
DR	MGI: MGI:103268; Vasp.
DR	INTERPRO: IPR000697; -.
DR	INTERPRO: IPR001960; -.
DR	INTERPRO: IPR002965; -.
DR	PFAM: PF005568; WHI: 1.
DR	PRINTS: PR01217; PRICEXTENSN.
SQ	SEQUENCE 375 AA; 39566 MW; 19362926CF4276C7 CRC64;

	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:	25:	26:	27:	28:	29:	30:	31:	32:	33:	34:	35:	36:	37:	38:	39:	40:	41:	42:	43:	44:	45:	46:	47:	48:	49:	50:	51:	52:	53:	54:	55:	56:	57:	58:	59:	60:	61:	62:	63:	64:	65:	66:	67:	68:	69:	70:	71:	72:	73:	74:	75:	76:	77:	78:	79:	80:	81:	82:	83:	84:	85:	86:	87:	88:	89:	90:	91:	92:	93:	94:	95:	96:	97:	98:	99:	100:
MS: Socia	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100			
MS: Socia	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100			
MS: Socia	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100			
MS: Socia	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100			
MS: Socia	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100			
MS: Socia	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	3																																																																			

DR INTERPRO; IPR001650; -.

DR INTERI

	RA	Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodis C.D., Krafcik C., Kravitz S., Kul D., Liao Z., Li X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu P., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.	
	RA	Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzy D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K.S., Nusakern D.R., Pacheb J.M., Palazzolo M., Pittman G.-S., Pan S., Pollard J., Puri V., Reese M.G., Relnert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skuski M.P., Smith T., Spier E.B., Spradling A.C., Stapleton M., Strong R., Sun E., Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A., Yeung J., Yen R.F., Zaveri J.-S., Zhang M., Zhang G., Zhao Q., Zheng L., Zhuo X.A., Ayres E.W., Dubin G.M., Venetog J.C., "Wenger"	
	RT	Sciences 287, 2195-2195(2000).	
	RX	EMBL; AF003581; RAF31170.1;	
	DR	FLYBASE: FBgn0016132; Chd1.	
	DR	INTERPRO: IPRO003330;	
	DR	INTERPRO: IPRO009533;	
	DR	INTERPRO: IPRO016501;	
	DR	INTERPRO: IPRO024641;	
	DR	PFAM: PF00176; SNF2.N.1.	
	DR	PFAM: PF00271; helicase_C.1.	
	DR	PFAM: PF00385; chromo.2.	
	DR	PROSITE: PS00598; CHROMO.1.1.	
	DR	PROSITE: PS00690; DEAD_ATP_HELICASE; UNKNOWN_1.	
	DR	PROSITE: PS50013; CHROMO.2.	
	SQ	SEQUENCE 1883 AA; 211826 MW; 2267ZCA899GF6E17 CRC64;	
		Query local BLAST results	
		Best Local Similarity    10.5%    Score 98.5; DB 5; Length 1883;	
		Mismatches 42; Conservative 24; Mismatches 73; Indels 43; Gaps 6;	
	OY	8 HSVTNLSQLGSPAGRGHRTLSSKASDADGSRS-----DSPLPHPKGKDSDFWARRE 60	
	DG	29 HSGSGSGSGS-----SCGSDSDSSGNSSGRSSPEPKSLSVACGPPTAA 76	
	OY	61 ASAALIGLDQKISHLTDEDLADFVDVKARVRGLVEDNESDGSEFFVFRDYMSNG 120	
	DG	77 AAAA----DKTNGFTDDFDSSSI-----SCGSDSDSDGAPGSDQR---NO 117	
	OY	121 TIDNFSLDTATMGTIPRSDGDGALEFGDKSKODANSVDVEELRQQQAOMEALKATK 180	
	DG	118 SINNAWTSSSLPPQNEDNET-----EXAQOPASDAADESSGSANVSPTSSTSS 172	
	OY	181 QE 182	
	DG	173 SE 174	
		RESULT 6	
	ID	OT4354 PRELIMINARY; PRT; 599 AA.	
	OC	OT4354	
	DT	01-NOV-1998 (TrEMBLrel. 08, Created)	
	DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)	
	DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)	
	DE	HYPOTHETICAL 66.0 KDa PROTEIN.	
	GN	SPBC25D12.02C.	
	OS	Schizosaccharomyces pombe (fission yeast).	
	OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;	
	OX	Schizosaccharomycetaceae; Schizosaccharomycetes.	
	NCBI_Taxid=4896;		
	RC	SEQUENCE FROM N.A.	
	RC	STRAIN=972H.	
	LA	Layne M.; Rajandream M.A.; Barrell B.G.; Whithread S.; Chillingworth T., Churcher C.M.;	

1

EMBL; AL031343; CAZ20830.1; -.  
 DR  
 EMBL; AF207839; AAF20943.1; -.  
 DR  
 INTERPRO; IPR000432; -.  
 DR  
 INTERPRO; IPR002863; -.  
 DR  
 Pfam; PF00488; Muts\_C.1.  
 DR  
 Pfam; PF01624; Muts\_N.1.  
 DR  
 PROSITE; PS00486; DNA\_MISMATCH\_REPAIR\_2; 1.  
 DR  
 PROSITE; PF001963; -.  
 DR

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

```
SQ SEQUENCE 1254 AA; 7A28EC8D8D2845B CRC64;

Query Match
Best Local Similarity 10.0%; Score 93.5; DB 3; Length 1254;
Matches 38; Conservative 33; Mismatches 71; Indels 43; Gaps 5;

QY 12 NSQLSPAPGRDHRTLSKAGSDADGRSDSLPHFRKGSDFYA----RRRASAALG 67
DB 121 HSLGSESPKLLRTSVKQPEDESEIDSPYKSKFKSLTSTQAEQDFPHYFYSKLEN 180
QY 68 LD-----OKISHLTDELDALFDQVQKARVRGLVEDNESD----- 105
DB 181 SELSEVDKPFIASSRRKKNVSAESDE-DEDFDAPATKGRSHKRIVSDSDDYVEPHI 239
QY 106 ----SQSSFPVRYKNSGTIDNFSLTDTATWPTPRSDGDLFFGKSKSQDAQND 161
DB 240 SSASSASLPIDVESMDVDYGSVSAAPIPK-----KESKSSNSL 287
QY 162 VEELR 166
DB 288 VESYR 292

RESULT 9
QY 421 PRELIMINARY; PRT; 3839 AA.
AC Q17632 PRELIMINARY; PRT; 1707 AA.
DB 31 IDKLSRL-DAELVYKQDKLRKREGPAKNYKQALVKQKMTQEQDMLAQSPNM 89
QY 113 RUKYMSGTIDNFSLTDTATWPTPRSDGDLFFGKSKSQDAQNDVYELRQQAQM 172
DB 90 E--QANTYTSLSK-DTKTTV-----DAMKLGKEMKAYKQVKYKIDQIEDLQDL 135
QY 173 EEALKTAKQ 181
DB 136 EDWEDANE 144

SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Rykatara G., Heves H.W., Mannhaupt G.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356834; CAB92704.1;
SQ SEQUENCE 3839 AA; 427187 MW; 6A9E57DEB48A5D9D CRC64;

Query Match
Best Local Similarity 9.8%; Score 91.5; DB 3; Length 3839;
Matches 47; Conservative 29; Mismatches 72; Indels 43; Gaps 9;

QY 3 OSLLRHVTSVNSQLG-SPAPGRDHRTLSKAGSD-----ADGDSRSDSLPHFRKGSDFM 54
DB 2338 KTLSEKTAVSMSQLGIPASG-----TAGTQAEDEFISGSSVSED-LTDDRETFDL 2389
QY 55 FYARRRASAAIL-GLQDKISHLTDELDALFDQVQKARVRG--LVEDNESDSQSSF 110
DB 2390 Y----RNSLTCLKLFGRDEFEDEDEDDYDDQYDDELDYGDMSQDNEINFSDEED 2446
QY 111 PVRCYKNSGTIDNFSLTDTATWPTP-----RSSDGDALFFGDKSKSQDAQNDVY 162
DB 2447 DLGEMGEGG-----MPSQPGVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 2494
QY 163 BELRQQAQME 173
DB 2495 QELSDDEDEDE 2505

RESULT 10
QY 421 PRELIMINARY; PRT; 219 AA.
ID Q9NZ23
```

```
AC Q9NZ23;
DB 01-OCT-2000 (TrEMBLrel. 15, Created)
DB 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DB 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HSPC177.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DB 121 NSQLSPAPGRDHRTLSKAGSDADGRSDSLPHFRKGSDFYA----RRRASAALG 67
DB 121 HSLGSESPKLLRTSVKQPEDESEIDSPYKSKFKSLTSTQAEQDFPHYFYSKLEN 180
QY 68 LD-----OKISHLTDELDALFDQVQKARVRGLVEDNESD----- 105
DB 181 SELSEVDKPFIASSRRKKNVSAESDE-DEDFDAPATKGRSHKRIVSDSDDYVEPHI 239
QY 106 ----SQSSFPVRYKNSGTIDNFSLTDTATWPTPRSDGDLFFGKSKSQDAQND 161
DB 240 SSASSASLPIDVESMDVDYGSVSAAPIPK-----KESKSSNSL 287
QY 162 VEELR 166
DB 288 VESYR 292

Query Match
Best Local Similarity 9.7%; Score 91; DB 4; Length 219;
Matches 34; Conservative 25; Mismatches 40; Indels 30; Gaps 6;

QY 66 LQKISHLTDELDALFDQVQKAR-----AVR-----RGLVEDNESDSQSSFPV 112
DB 31 IDKLSRL-DAELVYKQDKLRKREGPAKNYKQALVKQKMTQEQDMLAQSPNM 89
QY 113 RUKYMSGTIDNFSLTDTATWPTPRSDGDLFFGDKSKSQDAQNDVYELRQQAQM 172
DB 90 E--QANTYTSLSK-DTKTTV-----DAMKLGKEMKAYKQVKYKIDQIEDLQDL 135
QY 173 EEALKTAKQ 181
DB 136 EDWEDANE 144

RESULT 11
QY 421 PRELIMINARY; PRT; 1707 AA.
AC Q17632 PRELIMINARY; PRT; 1707 AA.
DB 31 IDKLSRL-DAELVYKQDKLRKREGPAKNYKQALVKQKMTQEQDMLAQSPNM 89
QY 113 RUKYMSGTIDNFSLTDTATWPTPRSDGDLFFGDKSKSQDAQNDVYELRQQAQM 172
DB 90 E--QANTYTSLSK-DTKTTV-----DAMKLGKEMKAYKQVKYKIDQIEDLQDL 135
QY 173 EEALKTAKQ 181
DB 136 EDWEDANE 144

SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
RA Rykatara G., Heves H.W., Mannhaupt G.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL356834; CAB92704.1;
SQ SEQUENCE 3839 AA; 427187 MW; 6A9E57DEB48A5D9D CRC64;

Query Match
Best Local Similarity 9.6%; Score 90.5; DB 5; Length 1707;
Matches 48; Conservative 30; Mismatches 62; Indels 79; Gaps 11;

QY 12 NSQLSPAPGRDHRTLSKAGSDADGRSDSLPHFRKGSDFM 55
DB 1501 NKKLKPFISGKYKRGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 1559
QY 56 YARREAAASAILGLQDKISHLTDELDALFDQVQKARVRGLVEDNESDSQSS----- 108
DB 1560 YLRKDSLIQA----KKKI---TENRRSL--EQKTPSIQIFNLMDQSSAASATEYSI 1610
QY 109 SFPV-----RDK-----YMSNG-----TIDN 124
```

```

Db 1611 SIPVPMVRQVNTSDKSDLSDEDDLTITREDAPPTISINLPRGRHALYSTIADIATEDD 1670
Qy 125 FSLDTAITMPTGTPRSDDDGDLFFGDKKSKQASNVDE 163
Db 1671 FYADSPVPMPTVPQADGS--FFGENDSRQYRQDSSE 1707

RESULT 12
Q41042 PRELIMINARY; PRT; 611 AA.
AC Q41042
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DE PROTEIN LOCALIZED IN THE NUCLEOLI.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
OC Fabales; Fabaceae; Papilionoideae; Pisum.
OX NCBI_TaxID=3889;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ALASKA; TISSUE=APICAL MERISTEM;
RA Tong C.G., Hsieh H.L., Blumenthal S., Reichler S., Balk J., Roux S.J.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L43510; AA74208.1; -.
DR HSSP; P09651; IUP.
DR MENDEL; 15263; Pisss; 2406; 15263.
DR INTERPRO; IPR000504; -.
DR PFAM; PF00076; rrm; 2.
SQ SEQUENCE 611 AA; 64774 MW; 7CFC749EF8ADEB64 CRC64;

Query Match 9.6%; Score 90; DB 10; Length 611;
Best Local Similarity 23.9%; Pred. No. 6.4;
Matches 45; Conservative 33; Mismatches 62; Indels 48; Gaps 9;

Qy 19 APCRHDRTLSKAGSDAGDSRSDSPLPHER-----GQSDMFYARREASAILGLDQKIS 73
Db 176 APAR--KAASSDEESDESDSEAFKAVVPAVKPA---AAKKKAESS-----DSDS 226
Qy 74 HLTDDELDA--LFDVOKARAVRGLVNEDES-----DSQSPFFVRDKYMSNG----- 120
Db 227 EDSDEEDTPTVYAKSVYAKRTEDDEDSDSDSDNDKTSKPSNKKRPVTV 286
Qy 121 -----TIDFSLDTAITMPTGTPRSDDDGDLFFGDKKSKQASNVDELRQQAOME 173
Db 287 SKDKKMNVDKUSS-----SDESDS-----DEESESEDPSTKPTOKTKDVEI 332
Qy 174 BALKTAQ 181
Db 333 DAKDSKK 340

RESULT 13
Q26189 PRELIMINARY; PRT; 757 AA.
AC Q26189
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DE MICRONEME PROTEIN-1 (FRAGMENT).
OS Plasmodium vivax;
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=PAPUA NEW GUINEA 32-1;
RA Cochrane A.H., Gladz B.W., Kamboj K.K.;
RL "Isolation of three distinct parasite populations expressing different
RT circumsporozoite proteins from the Philippine strain of Plasmodium
RT knowlesi.";
```

```

J. Clin. Microbiol. 26:178-181(1988).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=PAPUA NEW GUINEA 32-1;
RX MEDLINE-91004213; PubMed-2170017;
RA Adams J.H., Hudson D.E., Torii M., Ward G.E., Wellemans T.E., Aikawa M.,
RA Miller L.H.;
RT "The Duffy receptor family of Plasmodium knowlesi is located within
RT the microemes of invasive malaria merozoites.";
RL Cell 63:141-153(1990).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=PAPUA NEW GUINEA 32-1;
RX MEDLINE-91187056; PubMed-1849231;
RA Fang X.D., Kaslow D.C., Adams J.H., Miller L.H.;
RT "Cloning of the Plasmodium vivax Duffy receptor.";
RL Mol. Biochem. Parasitol. 44:125-132(1991).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=PAPUA NEW GUINEA 32-1;
RX MEDLINE-92357776; PubMed-1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=PAPUA NEW GUINEA 32-1;
RA Tsuboi T., al-Yaman F., Prickett M.D., Alpers M.P., Adams J.H.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE OF 1-751 FROM N.A.
RC STRAIN=PAPUA NEW GUINEA 32-1;
RX MEDLINE-95048797; PubMed-7960140;
RA Tsuboi T., Kappe S.H., al-Yaman F., Prickett M.D., Alpers M.,
RA Adams J.H.;
RT "Natural variation within the principal adhesion domain of the
RT Plasmodium vivax Duffy binding protein.";
RL Infect. Immun. 62:5581-5586(1994).
DR EMBL; L23072; AAA61769.1; -.
FT NON_TER 1 757
FT NON_TER 757 757
SQ SEQUENCE 757 AA; 84006 MW; A3B354E25F212D3C CRC64;

Query Match 9.6%; Score 90; DB 5; Length 757;
Best Local Similarity 25.7%; Pred. No. 8.2;
Matches 47; Conservative 24; Mismatches 72; Indels 40; Gaps 8;

Qy 6 LRHVTNSQLGSPAPGRHRTLSKAGSDAGDSRSDSPLPHERGKDSWDFYARREASAI 65
Db 429 LRHSKDNDS-----GPAEFAESTKSAESMANPDSNKGKGTG--KGQDNDAKATKDSNSS 482
Qy 66 LGDQKISHLTDELDAFDVOKARAVRGLVNEDES-----DSQSPFFVRDKYMSN--GTI 122
Db 483 DG-----TSSATGDTTDAV-----DREINKVPED-----DRTVSGKDGGE 520
Qy 123 DNFSLDTAITMPTGTPRSDDDGDLFFGDKKSKQASNVDEE---LRQQAOMEALKA 179
Db 521 DNSAKNDATVVGEDRTRENSAG-----GSTNDRSKNDTEKNGASTPDSKQSEDATLS 574
Qy 180 KQE 182
Db 575 KTE 577

RESULT 14
Q60465 PRELIMINARY; PRT; 590 AA.
AC Q60465
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TEMBLrel. 08, Last annotation update)
DE PEST REPEATS ACIDIC PROTEIN GADD34.
```





